International Journal of Electrical and Computer Engineering (IJECE)
Vol. 9, No. 6, December 2019, pp. 5192~5204
ISSN: 2088-8708, DOI: 10.11591/ijece.v9i6.pp5192-5204

Robust feature extraction methods for general fish classification

Mutasem K. Alsmadi[1], Mohammed Tayfour[2], Raed A. Alkhasawneh[3], Usama Badawi[4], Ibrahim Almarashdeh[5], Firas Haddad[6]
[1,2,4,5]Department of Management Information Systems, College of Applied Studies and Community Service, Imam Abdulrahman Bin Faisal University, Saudi Arabia
[3,6]Department of General Courses, College of Applied Studies and Community Service, Imam Abdulrahman Bin Faisal University, Saudi Arabia

Article Info

Article history:

Received Mar 2, 2019
Revised Jul 6, 2019
Accepted Jul 17, 2019

Keywords:

Anchor points measurements
Back propagation algorithm
Meta-heuristic algorithm
Features extraction
Statistical measurements
Texture measurements

ABSTRACT

Image recognition process could be plagued by many problems including noise, overlap, distortion, errors in the outcomes of segmentation, and impediment of objects within the image. Based on feature selection and combination theory between major extracted features, this study attempts to establish a system that could recognize fish object within the image utilizing texture, anchor points, and statistical measurements. Then, a generic fish classification is executed with the application of an innovative classification evaluation through a meta-heuristic algorithm known as Memetic Algorithm (Genetic Algorithm with Simulated Annealing) with back-propagation algorithm (MA-B Classifier). Here, images of dangerous and non-dangerous fish are recognized. Images of dangerous fish are further recognized as Predatory or Poison fish family, whereas families of non-dangerous fish are classified into garden and food family. A total of 24 fish families were used in testing the proposed prototype, whereby each family encompasses different number of species. The process of classification was successfully undertaken by the proposed prototype, whereby 400 distinct fish images were used in the experimental tests. Of these fish images, 250 were used for training phase while 150 were used for testing phase. The back-propagation algorithm and the proposed MA-B Classifier produced a general accuracy recognition rate of 82.25 and 90% respectively.

Copyright © 2019 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Mutasem Alsmadi,
Department of MIS, College of Applied Studies and Community Service,
Imam Abdulrahman Bin Faisal University,
Al-Dammam, Saudi Arabia.
Email: mksalsmadi@gmail.com, mkalsmadi@iau.edu.sa

1. INTRODUCTION

The traditional image recognition process mostly employed the skills and senses of human which has caused inaccurate and unsatisfactory recognition process. Due to the importance of this process, computers are now used for recognition processes due to their level of accuracy and efficiency. In this regard, a number of approaches have been employed for image processing [1-8] and pattern recognition [9-11]. Image recognition has been actively studied and several researches have been conducted on this subject. There are in fact many issues associated with image recognition such as distortion, object overlap and blockade in digital images, and also errors in the results of segmentation [12-15]. As has been shown in present researches, the already available fish recognition systems are still lacking in certain areas. For instance, the current systems are still not sufficiently able to detect and classify fish. Apart from that, significant number of daily deaths still happens owing to failure in making distinction between fishes that are dangerous and those that are not [16-20].

The main contribution of this work is to develop a system for fish image classification, where this system employs anchor points, texture as well as statistical measurements for features extraction. Accordingly, the classification of fish images has been chosen as the focal point of this study. Fish image of certain size and format is the input of the proposed system, and using the anchor points, texture and statistical measurements, features of the fish images will be extracted. Then, using the MA-B Classifier on the extracted features, the fish in question would be classified into dangerous and non-dangerous fish. This fish image will further be classified as Predatory or Poisonous fish (if dangerous), or garden or food fish (if non-dangerous). This study would be of value to many arenas including marine, industry and agriculture.

The rest of the paper is organized as follows: Section 2 discussed the literature review. Section 3 describes materials and methods employed. Section 4 describes the classifier architecture. Section 5 discusses and evaluates the obtained results using the proposed MA-B Classifier and other classification methods. Section 6 concludes the paper.

2. LITERATURE REVIEW

Fish recognition is very complicated and difficult task but is useful to business and agriculture. Distortion, overlap, noise, distortion, occlusion, and also error in segmentation are among the challenges faced in achieving accurate and reliable fish recognition [21-34]. Among the works relevant to this study is one from Mokti and Salam [35]. In their work, the authors applied a hybrid of Mean-shift and median-cut algorithms in their segmentation of color fish images. Prior to that, the image pre-processing technique was applied in order to improve the image before its color space was transformed into LUV color space. Then, to cluster around the region, the authors applied the mean-shift segmentation. However, since the segmentation executed by Mean-shift algorithm is of low level, certain region carries no semantic meaning. Hence, median-cut algorithm can be used as solution to this problem. Furthermore, this algorithm has the capability to reduce the color depth inside the image. With the application of Mean-shift and median-cut hybrid, the results were improved especially in terms region grouping. Besides that, the authors were able to achieve clearer borders of segmented regions easier. Using the proposed algorithm provides the advantages of Mean-shift segmentation method. At the same time, the weaknesses associated with the use of median-cut grouping method are evaded.

In Alsmadi *et al*. [36], a fish classification prototype was proposed. This prototype combines between the features extracted from measurements of shape and size using the measurements of distance and geometry. The authors employed 20 different families of fish containing different fish types each, and the sample used contains a total of 350 different types of fish images. These images were split into two datasets, with 257 training images in one dataset and 93 testing images in the other. The authors attained 86% accuracy when using the neural network associated with the back-propagation algorithm on the dataset of used test. From the results, the authors proved the ability of the proposed classifier in categorizing the fish into its correct class, in categorizing the classed fish into poison or non-poison fish, and in categorizing the poison and non-poison fish into its correct family.

In Alsmadi *et al*. [18], the extracted features from color texture measurements was used in combination. In particular, the authors employed gray level co-occurrence matrix (GLCM) for the production of a prototype for classifying fish. In testing the prototype, the authors used 20 different fish families containing different number of fish types each. Altogether, there were 610 different fish images were used. These images were classed into two datasets. One dataset contains 500 training images while the other contains 110 testing images. Neural network connected to the back-propagation algorithm was used in this work, and the authors achieved 84% accuracy on the test dataset. With the application of the proposed classifier, the authors were able to categorize the fish into its cluster, categorize the clustered fish into poison or non-poison ones, and further categorize the poison and non-poison fish into its matching family.

In Alsmadi *et al*. [17] a fish classification prototype was presented. This prototype combines features extracted from color signature measurements. Here, histogram of color, RGB color space, in addition to GLCM was used. Accordingly, the authors used a crop out of color signature for differing families of fish. The poison and non-poison fish were all used for the extracted color signature features. The authors used 20 different families of fish in testing the proposed system. In each family, there were different types of fish. Overall, there were 610 different images of fish in the sample used and these images were divided into two datasets: 400 training images in one dataset and 210 testing images in the other dataset. The authors used neural network associated with the back-propagation algorithm and it yielded 84% overall accuracy on the employed test dataset. This study proves that the classifier it proposed allows fish to be classed into its corresponding cluster. Furthermore, the clustered fish could further be classed into poison or non-poison ones, and into its appropriate family.

Badawi and Alsmadi [37] demonstrated a generic fish classification system. This system employs a hybrid metaheuristic algorithm, which is genetic algorithm with iterated local search, in addition to back-propagation algorithm (GAILS-BPC). Classification was executed based on a combination between significant extracted features with the use of anchor points and texture and statistical measurements; this allows fish images to be classed into dangerous and non-dangerous families, while dangerous families of fish can be classed into Predatory and Poison fish family. Furthermore, families of non-dangerous fish can be classed into garden and food fish family. This study used 24 fish families in testing the prototype, and the families of fish employed in this work contain different number of species of fish each. There were two phases, namely training phase and testing phase. Training phase employed 220 fish images while testing phase employed 100 fish images. Hence, overall, this study used 320 different fish images. The overall accuracy of recognition rate obtained by the authors when using the prototype with GAILS-BPC was 80.5%.

In Alsmadi *et al*. [9], a prototype that uses Hybrid Memetic Algorithm (Genetic Algorithm and Great Deluge Local Search) and Back-Propagation Classifier (HGAGD-BPC) and Back-Propagation Classifier (BPC) for classifying fish was demonstrated. Here, the authors performed the classification task following the combination between the extracted features obtained from Potential Local Geometric Features (PLGF) and Shape Features. In this work, HGAGD-BPC achieved better results at 96%. Somehow, compared to BPC, HGAGD-BPC had high computational time, but BPC showed lower percentage at 86%. These classifiers allow fish to be grouped into its corresponding cluster. Then, the clustered fish could be further classed into poison or non-poison fish, and into its rightful family.

Sayed *et al*. [38] demonstrated an automated fish species identification system following a modified crow search optimization algorithm in their work. Median filtering was used to generate smooth image and eliminate the noise. This reduces the diversity of intensities between the neighbors. A k-mean clustering algorithm was then used to segment the fish image into multiple segments, which brought forth the feature extraction process based on shape and texture for the task of classification. The data dimensionality of the extracted features was decreased in this work, and for the purpose, a new modified binary version of crow search algorithm was applied. The classification task involved the use of support vector machine and decision trees and the species of fish were classed following either their class (e.g., Actinopterygii and Chondrichthyes) or their order. The authors worked on 270 images in different species, classes and orders on the proposed system and reported the superiority of the proposed system compared to other advanced algorithms. The authors also reported that the overall fish species system of identification yields 10 folds on average, 96% accuracy of classification for classification based class and 74% for classification based on fish order.

3. RESEARCH MARERIALS AND METHODS

3.1. Dataset

A total of 400 fish images were used in this study. The images were from Global Information System (GIS) on Fishes (fish-base), and they were obtained in September, 2013. The images included real-world images of fish captured in "controlled", "out-of-the-water" and "in-situ" settings. For the "controlled" images, they encompass fish specimens in position where their fins are spread and the images were captured in consistent background with illumination that is controlled. The dataset employed in this work contains three categories; each of them contains different number of fish families as following:

a. Dangerous Fish Families: Carcharhinus Leucas, Carcharodon Carcharia, Atractosteus Spatula and Hydrocynus Goliath.
b. Poison Fish Families: Red Snapper, Trigger, Porcupine and Thorn.
c. Garden and Food Families: Acestrorhynchidae, Acropomaatidae, Albulidae, Anomalopidae, Caesionidae, Drepanidae, Istiophoridae, Leiognathidae, Megalopidae, Platycephalidae, Priacanthidae, Scombridae, Siganidae, Sillaginidae, Stromateidae and Triacanthidae.

3.2. Texture features calculating using GLCM

The computation of texture features according to GLCM involves seven steps. Image acquisition is the first and most important step. The second step involves the transformation of digital fish image into gray scale image whereby digital fish images are the ones commonly dealt with in this research. Notably, fish differs in terms of shape, and therefore, a crop faction is employed in order to manually determine the fish shape so that error could be eradicated. Then in the third step, a crop out of the pattern of interest, in this case, it is the shape of the fish shape, is separated from the background. Using this approach, high quality fish recognition can be attained.

In the fourth step, the captured crop is filtrated from fish pixels. For the purpose, a 5×5 Gaussian Filter is used. The filtrated crop image is next split into blocks with the size of 4×4; this is the fifth step. This is followed by computation of image quality features for each block in accordance with the GLCM (this is the sixth step). Lastly, in the seventh step, the attained features are stored. In brief, the computation of fish's texture features by GLCM is as outlined below:

Step 1: digital fish image acquirement.
Step 2: conversion of image to gray scale image.
Step 3: manual determination of fish shape using a crop faction to eradicate errors, while the crop out of the pattern of interest (fish shape) is subtracted from the background.
Step 4: filtration of captured crop out of fish pixels with a 5×5 Gaussian Filter.
Step 5: division of the filtrated crop image into 4×4 blocks.
Step 6: Calculation of image quality features for all blocks using GLCM. Computation on 24 discrete image quality features for the "filtrated crop" following the four directions (horizontally (90^0), vertically (0^0), and two diagonally (45^0 and 135^0)), specifically, Average or mean value, Standard Deviation, Contrast, Dissimilarity, Homogeneity, and Energy.
Step 7: storage of the attained features in the database.

3.3. Shape features calculating using anchor points location detection

In measuring fish shape, several anchor points need to be identified, as shown in Figure 1. Detection of anchor points has indeed been the interest of many since past several years particularly among those working in pattern recognition. Points' detection is used to find a significant set of points which will facilitate the attainment of anchor measurements for patterns of interest, in this case, fish object. In this study, anchor point detection is used for the determination of 23 labeled points which will facilitate the determination of the location of each feature in fish image recognition. This is followed by the computation of the geometrical features with the use of the determined anchor points; this is for classifying the fish. When the entire anchor points over the fish object have been detected, distance and angle measurements are used to extract the significant features.

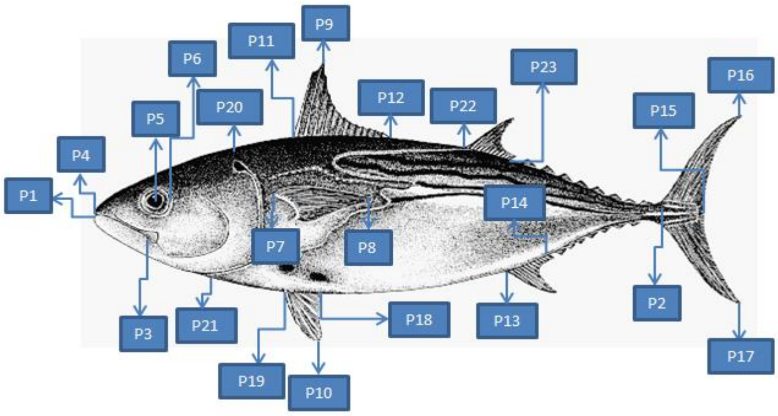

Figure 1. The locations of the anchor point measurements

Shape measurements involve the computation of the edge and distance measurements of the fish object as well as the determination of the significant identical and differing parts for each family of fish. Furthermore, classification of greater accuracy would result if the procedure of classification employs the measurements of vector's angles employing three points for each angle of the caudal fin angle and fish head angle [36]. In addition, application of distance measurements allows the determination as well as extraction of several features including the radius of fish eye and length of pectoral fin.

3.4. Distance and angle measuring tools

Distance and angle measurements are used to compute shape features. In detail, the distance measurements comprise the distance between 21 anchor points namely: P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12, P13, P14, P15, P18, P19, P20, P21, P22, and P23. Table 1. Twenty two extracted features from the determined anchor points. Figure 1 provides the details. Meanwhile, the angle between three anchor points over the fish object makes up the angle measurements such as: eye-end mouth angle, Caudal fin angle and fish head angle. Figure 1 provides the details. The anchor points and the feature selected are computed utilizing distance and angle measurements. These anchor points and feature are presented in Table 1 and Table 2, and their details are provided in the ensuing subsection.

Table 1. Twenty two features calculated using the determined distance measurement

Distance. #	Anchor Points Features	Distance. #	Anchor Points Features
D1	(P1 and P2)	D12	(P13 and P14)
D2	(P11 and P19)	D13	(P2 and P15)
D3	(P1 and P3)	D14	(P19 and P10)
D4	(P3 and P5)	D15	(P20 and P21)
D5	(P5 and P6)	D16	(P22 and P23)
D6	(P7 and P8)	D17	(P5 and P3)
D7	(P11 and P12)	D18	(P13 and P2)
D8	(P12 and P22)	D19	(P4 and P3)
D9	(P18 and P13)	D20	(P4 and P11)
D10	(P2 and P5)	D21	(P5 and P4)
D11	(P5 and P20)	D22	(P19 and P4)

Distance measurements are integral in pattern recognition particularly in the task of robust features extraction to improve the accuracy of classification. In algebraic geometry, the computation of the distance ‘D’ between the points C= (a1, b1) and E = (a2, b2) is expressed in formula 1.

$$D = \sqrt{(\Delta a)^2 + (\Delta b)^2} = \sqrt{(a_2 - a_1)^2 +}\ (b_2 - b_1)^2 \quad (1)$$

As shown in Figure 1, there are 23 anchor points, and these points denote the length between anchor points as Table 1 is showing. Hence, with the application of the distance measurement formula, a total of 15 features were attained. The angle between two vectors, suspended by one point, is dubbed the shortest angle at which one of the vectors has to be rotated to the position that is co-directional with another vector [39]. The formula below computes the angle θ between two vectors:

$$\cos a = \frac{|\bar{a}.\bar{b}|}{|\bar{a}|.|\bar{b}|} \quad (2)$$

The resultant eight features that computed with the angle measurements according the anchor points highlighted in Figure 1 are shown in Table 2.

Table 2. Sixteen features that were calculated using the determined anchor points

Angle. #	Anchor Points Features	Angle. #	Anchor Points Features
A1	(P15, P9 and P16)	A9	(P16, P4 and P17)
A2	(P9, P4 and P10)	A10	(P9, P15 and P10)
A3	(P9, P17 and P16)	A11	(P1, P16 and P17)
A4	(P21, P16 and P17)	A12	(P20, P16 and P17)
A5	(P16, P15 and P17)	A13	(P23, P13 and P1)
A6	(P20, P4 and P21)	A14	(P23, P13 and P115)
A7	(P5, P3 and P16)	A15	(P9, P16 and P10)
A8	(P5, P3 and P17)	A16	(P15, P10 and P17)

3.5. Statistical measurements

Statistical measurements were carried out utilizing the features extracted from images of fish belonging to 24 fish families. This is will determine and ascertain the significant features which will assist in the attainment of recognition of accuracy while also recognizing the fish images into either dangerous or non-dangerous family. Accordingly, the outcomes of correlation grounded upon the features extracted with the application of measurements of anchor points can be viewed in Figure 2. The statistical outcomes from the extracted features demonstrate the diverse correlation value between certain extracted features particularly head, eye and caudal angles of the fish. These features can in fact be deemed as good features for improving the accuracy of classification, which, are of value in the context of this work. As an example, the dangerous fish families show negative correlation value between the head and eye angles. Described differently, for fish belonging to these families; the head angle increases as the eye angle decreases. Furthermore, the value of correlation between the caudal and eye angle for fish in these families also appears negative.

On the other hand, fish belonging to certain non-dangerous fish families appear to have positive correlation value between the head and eye angle. This means that for fish belonging to these families; the head angle increases as the eye angle increases. Additionally, the values of correlation between the caudal and eye angles also appear to be positive. As can be construed from the findings, the resultant values of correlation of the extracted features differ based on family. The correlation values will increase the distinction between these families of fish (poison, non-poison, wild and food fish families).

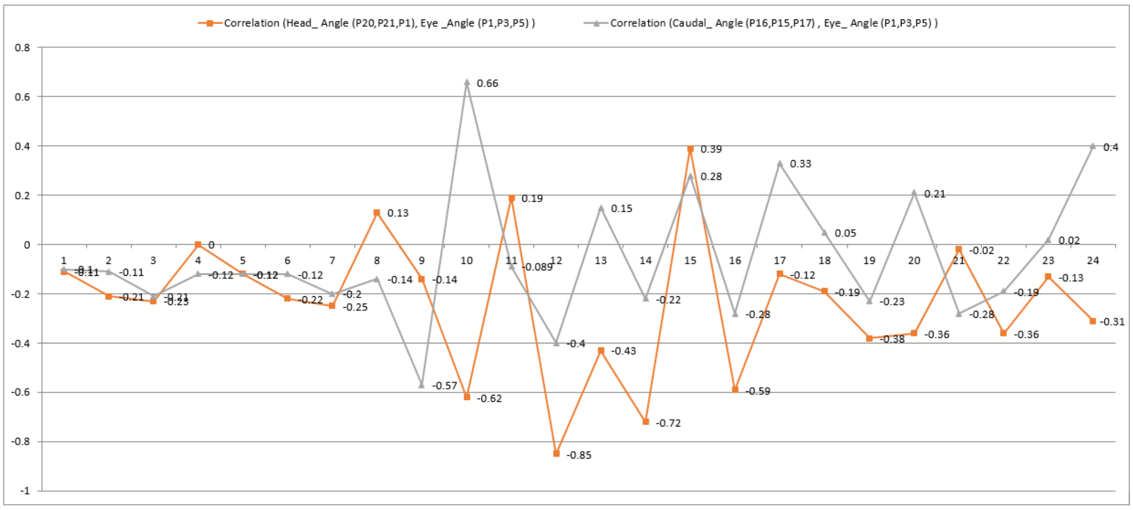

Figure 2. Correlation results based on the features that were extracted using anchor points measurements

4. CLASSIFIER ARCHITECTURE

4.1. Genetic algorithm

A genetic algorithm (GA) encompasses a heuristic approach and it is grounded upon population which simulates the natural selection procedure. GA creates valuable novel solutions to challenging problem using sample solutions in a population. There are three key phases in GA whereby the first phase concerns the selection technique to select two solutions from the population and then recombine them. Relevantly, several techniques of selection were proposed in Michalewicz [16] including Tournament Selection, Truncation Selection and Roulette Wheel Selection. The second phase includes the use of a crossover operator to carry out the mating process. Crossover is the genetic way to discover novel solutions, i.e., solutions with better value of fitness, within the search space. The third phase involves the use of a Mutation operator for discovering the neighbor solutions. Mutation operator is a local search. During this phase, the population is updated. Furthermore, the generation of solutions with better value of fitness will improve the quality of search space [12].

4.1.1. Initialization

The production of weights for feed-forward artificial neural network is usually a random process. Nonetheless, the representation of chromosome greatly contributes to the success of genetic algorithm. In this research, a simple presentation for chromosome grounded upon binary representation is applied for each solution. Here, the chromosome denotes several weights (real values) that are randomly obtained from the matrices weight. These weights are in fraction numbers form, these fraction numbers are signified in the genes within the chromosome in binary strings.

4.1.2. Roulette wheel selection

Roulette Wheel Selection is the most commonly used method. Roulette Wheel Selection was the creation of Baker (1987) and it is known as the simplest schema of selection. As explained in Leung *et al*. (2003), in the application Roulette Wheel Selection, two chromosomes from the population are chosen to go through genetic operations for reproduction using the spinning the roulette wheel method. Parents with high potential arguably will generate better offspring, which have better chance of survival. The chromosome with higher value of fitness is likely to have greater opportunity in being chosen as parent.

During the selection process, stochastic selection is made from one generation, and the selection becomes the foundation for the ensuing generation. The rule of thumb is that the fittest ones have greater survival opportunity as opposed to the weaker ones, just like the real life situation in nature. The fitter ones will then form the mating pool for the following generation. However, the weaker ones still have a chance as they may carry valuable genetic coding for the next generations.

4.1.3. Crossover

Crossover is the core of genetic algorithm, and it denotes a mating process. The process of crossover aims at discovering new solutions within the search space. In a customary crossover operator, the pairing of individuals of the population is done in a random manner. Two mating chromosomes are cut one time at the respective points and the sections following the cuts are exchanged. The point of crossover usually can be randomly selected. Described in more detail, the recombination of individuals generates new individuals by having the information from two or more parents combined. This is generally executed through the merging of parents' variable values. Single point crossover is regarded as the simplest crossover method. As described in Ansari and Hou (1999), the outcome of this method is one or two child string through the random selection of crossover location within pattern string length. In the execution of single point crossover, a point is randomly selected. Then, the parents' chromosomes will be severed from that point, and exchange will be made to the resultant sub-chromosomes.

4.1.4. Mutation

In genetic algorithm, mutation functions as a local search for the discovery of the neighbor solutions. Mutation comprises a genetic operator for sustaining the diversity of genes from one generation of a population of chromosomes to the forthcoming one. Following the process of crossover, the resultant individuals or weights will go through the process of mutation. Mutation usually replaces the genes that are lost during the evolutionary process in a different form or generates new genes that were not explored in the original population. With mutation, the algorithm could prevent from being trapped in local minima. This is because the population of chromosomes will be prevented from becoming too identical to one another which would cause a slow down or a halt to the evolution. Accordingly, a variable is chosen with specified probability, while its value is altered using a random value. A non-uniform mutation method is applied in this research. This method transforms one of the genes belonging to the parent in accordance with the non-uniform probability distribution.

4.1.5. Fitness function

The fitness function provides assessment on the performance of each individual, and it is based on problem. Here, the performance of each individual is computed using the percentage Variance Account Function (VAF) between two signals. As highlighted in Sheta (2006), the computation of VAF follows the formula below:

$$V = 1 - \frac{\text{variance(y−y\_est)}}{\text{variance(y)}} * 100\% \quad (3)$$

Based on the expression above: y denotes the real output; y_est denotes the projected output of a model, and VAF is quantified for the two signals to generate the output V. For the two signals, the VAF is equivalent to 100%, and should they have different value, the VAF will be less. When y and y_est have more than one column, the computation of VAF is made for each column in y and y_est. In general, VAF is applied in the verification of the model's accurateness through the comparison of the real output with the model's projected output.

4.1.6. Stopping criterion

In GAs, a stopping criterion is generally signified by the maximum number of generations. Nonetheless, when ideal value of fitness (i.e. optimal weight) can be achieved, a stopping criterion is also considered to have been achieved. However, in this research, maximum number of generations is used with no consideration on whether the ideal fitness value is attained or not.

4.2. Simulated annealing algorithm

Simulating Annealing (SA) introduced by Kirkpatrick [40] has greater robustness as opposed to simple local search owing to the fact that it also accepts worse solutions with some probability [41, 42]. SA has been popularly employed in the solution of hard combinatory problems. Apart from that, the use of SA attempts to prevent entrapment in local optimum solution through the allotment of probabilities to moves that appear to be deteriorating. In this regard, SA could accept solutions, which, as opposed to past ones, are neither better nor much worse, which allows escape from local optimum and discovery of the global one [43, 44]. The authors in [41] provided a generic SA algorithm for problem of maximization. SA begins with best weight solution (s) picked out from the pool of population within the genetic algorithm. The beginning temperature is at 1000 when the search begins and the final temperature is at 0 with iteration number, #Iter is set to 10000. At each iteration, the temperature is randomly reduced through the creation of a fraction number between 0 and 1. Next, a neighbor is outlined by the indiscriminate creation of a random fraction number between 0 and 1, which is included into the solution value. Applying the formula expressed in Equation (3), computation is made on the fitness function value of the new neighbor. A worse solution is received if the indiscriminately produced number is lower than e -δ/T, where δ = f(s')-f(s), T=ck (ck denotes the temperature of the present iteration number). Then, update is made on the present solution (s). This process ensues until the highest number of iterations is achieved, i.e., 1000 iterations.

4.3. Neural network model

Neural network with BP algorithm is used for training and classification purpose [45]. The selection of the neurons number for the input and hidden layer was grounded upon the experiment performed in this study. This allows the decision on the appropriate number of neurons for the improvement of accuracy of the classification [46]. Meanwhile, as MA-B Classifier will classify 24 fish families; there will be 24 neurons within the output layer.

4.4. Memetic algorithm

Hybridization of two or more algorithms together has been attempted by countless scholars for the purpose of improving the performance of the search algorithms [47, 48]. Such attempt is underpinned by the notion that hybridization allows the merging of the best features from one algorithm with those of others [49, 50]. As highlighted in Moscato in [51], memetic algorithm is an augmentation of genetic algorithm, except that here, a local search is applied on individuals following genetic operators for instance, simulating annealing, and steepest descent algorithm.

Table 3. Neurons number for each neural network layer

Classifiers	NO. Neurons in layers		
	Input .Layer	H. Layer #1	Output. Layer #3
BP	64	25	24
MA-B	64	25	24

Table 4. Memetic algorithm parameters setting

Parameter	Value
Iterations number of GA	1000
Rate of Crossover	0.09
Rate of Mutation	0.02
Initial temperature	1000
Final temperature	0
SA generation number	600

According to Tan *et al*. in [52], the use of local search algorithm improves the exploitation process not exploration process. Relevantly, among countless scholars, memetic algorithms have been applied to improve the standard genetic algorithm performance-wise. Nonetheless, among a number of researchers the term "hybrid" is used instead of the term "memetic" when genetic algorithm is combined with local search approach. Accordingly, the number of neurons for each neural network layer is presented in Table 3. Table 4 shows the parameters setting of the of memetic algorithm.

5. EXPERIMENTAL RESULTS AND DISCUSSION

In previous studies such as [53] performed fish recognition based on texture features, the texture features were extracted only from the fish ventral part, This limited area (ventral part) reduces the effectiveness of the extraction of texture features in the classification stage, because the values and relationships between the neighboring pixels of the fish ventral part texture are converged to each other, where this makes it difficult for the classifier to accurately recognize the processed fish image [18, 26, 54]. The tail of fish is also taken into account in this context. Some fishes have forked tail while others have rounded tail. The tail shape nature is considered in this study. According to [55], some fishes have double emarginated shape tail, while some have lunate shape or forked shape tail. The spaces that are present between the shape characters of fish object are very much affected by the extracted texture feature values according to texture measurements. Briefly explained, the differences texture between the background image and the fish object comprise the distinction between the fish families, particularly between poison and non-poison fish family according to texture measurements.

Therefore; this work intended to recover this limitation by utilizing the differences in texture between the background image and the fish object . In certain instances, certain fishes possess dorsal fin and adipose fin and there is a gap (inter dorsal-adipose space) between these two types of fin. On the other hand, among some identical fishes, their dorsal fin and adipose fin are close together with no gap. Table 5 shows the difference between the extracted shape features from poison fish family (Red Snapper) and non-poison fish family (Scombridae). As can be seen in Table 5, in certain situations, some poison fish have smaller distance measurements as opposed to non-poison fish. Suppose that for poison fish object, the length of its mouth is 15.3 pixels and that of its dorsal fin is 23.66 pixels. Meanwhile, for the Scombridae fish, the length of its mouth is 25.33 pixels and that of its dorsal fin is 66.77 pixels. In the measurements of angle, for the Scombridae fish, the caudal fin angle is 116 pixels, and for the Red Snapper fish, the caudal fin angle is 145 pixels. In the measurements of shape, for the Scombridae fish object, its contour of length is 1643 pixels while that for Red Snapper fish object, it is 2625 pixels. Notably, among some families, they have similar shape characters.

Table 5. Some extracted shape features from both poison and non-poison fish families

Shape Features	Non-Poison Fish	Poison Fish
Fish Mouth length (P1 and P3)	25.33	15.3
Distance between the right-end of mouth and the start of dorsal fin (P3 and P11)	177	104
Dorsal Fin Length (P11 and P12)	66.77	23.66
Caudal Fin Angle (P16, P15 and P17)	116	145
Length contour of fish	1643	2625

However, each family carries its own traits that are specific to species. For example, forked tails are a trait to many species. As such, it is necessary to employ multiple shape characteristics in making classification to many differing families. Furthermore, as indicated in [56, 57], the use of the characteristics individually may not offer complete identification. Rather, the characteristics need to be used in combination in order to provide adequate information in the classification of numerous families into their corresponding families. Twenty four features were extracted using GLCM method utilizing gray texture measurements, 22 distance features were extracted utilizing distance measurements, 16 angle features were extracted utilizing angle measurements, and 2 features were extracted using statistical measurements. The back-propagation classifier was applied using a set of input features, but there are issues associated with this classifier including entrapment in the local optima and low rate of convergence [9]. As a solution, a hybrid meta-heuristic algorithm (MA-B Classifier) was proposed in this study. The meta-heuristic algorithm solves the optimization problem. Furthermore, as opposed to the conventional back-propagation algorithm, meta-heuristic algorithm demonstrates high level of effectiveness in the prevention of getting trapped in the local optima.

Accordingly, BP and MA-B Classifiers were used in testing the extracted features. As such, the achieved outcomes demonstrate the success of the features extraction and recognition methods in achieving high classification accuracy as opposed to past methods. In fact, the best accuracy results were at 87% and 95%, while the worst ones were at 81% and 87% respectively as shown in Figures 3 and 4.

The results show variation and this is attributed to identicalness of shape and texture of most fish families with one another. The original pixel values may also be present and this causes identical values of extracted features which will cause the complexity of the extracted features to increase. MA-B Classifier proposed in this work will train and classify these features. Notably, some families of fish carry their own species-specific-traits, and this facilitates MA-B Classifier in classifying them. For instance, some of the non-poison fish family has angle of upper triangle resembling other dangerous fish families. Also, these non-poison fishes carry some distinctive features including the space length between the right-end of first dorsal fin and the beginning of second dorsal fin, Pelvic fin length and Head width. Figures 3 and 4 present the outcomes of recognition accuracy for each fish family. Hence, this study was able to successfully recognize the families of dangerous fish with high classification accuracy owing to their species-specific traits (different shape compared with other family) that are distinct from other non-poison and poison families of fish.

As opposed to the conventional methods, the extracted features using the proposed methods (anchor points, texture and statistical measurements) with the proposed BP and MA-B Classifiers show superior performance over the state of the art methods used in [34, 37] particularly with respect to recognition accuracy with a percentage of 82.25 and 90 as shown in Figure 5. The methods of anchor points and texture measurements appear to be less influenced by the expression of fish and the global variations in fish object presence within the image. Furthermore, MA-B Classifier shows better performance when compared to the conventional BP classifier according to the features extracted with GLCM, angle as well as distance measurements. Using SA with GA, the recognition accuracy of MA-B Classifier is substantially improved as the weights to be used in the process of BPC training and processing are improved and optimized.

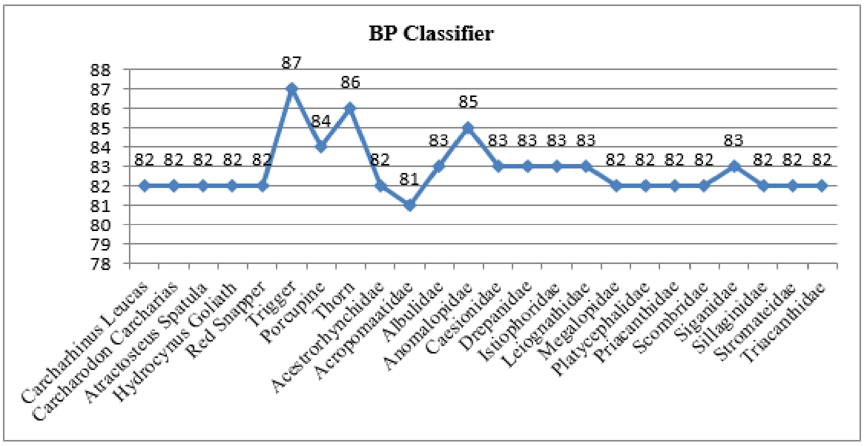

Figure 3. Recognition accuracy results using BP classifier

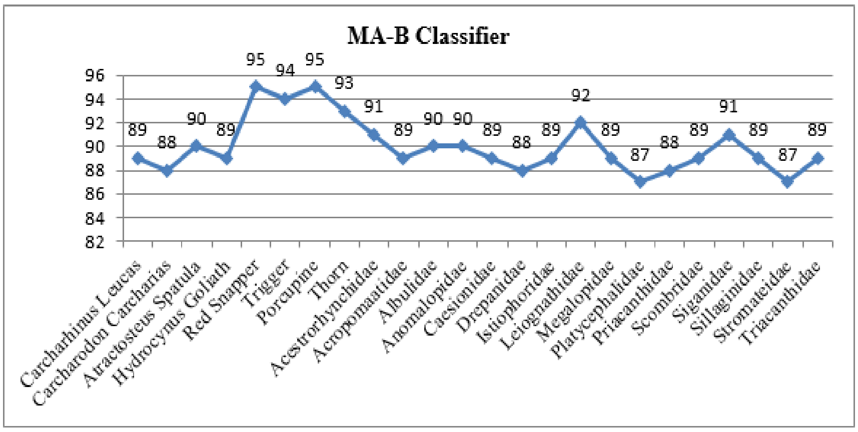

Figure 4. Recognition accuracy results using MA-B classifier

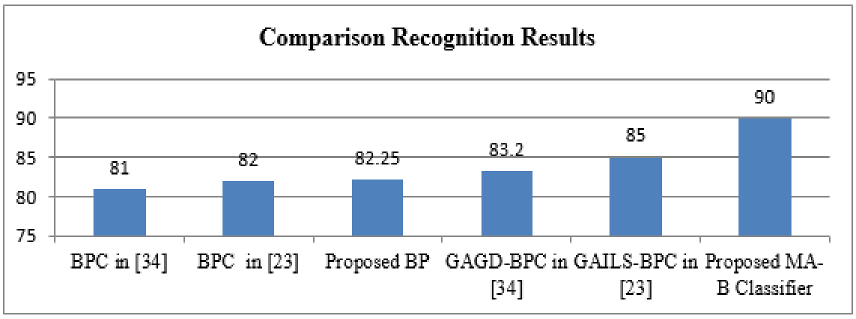

Figure 5. Comparison graph for overall accuracy results between proposed BP and MA-B classifiers and other comparative methods

6. CONCLUSION

This study constructed an innovative computer vision system that allows automatic recognition of invasive or other unknown species, using their phenotypes. The features extraction methods are grounded upon noteworthy merged features obtained with GLCM, anchor points detection, and statistical measurements from texture and shape measurements. GLCM was used to extract 24 features while angle and distance tools were used to extract 39 features whereas statistical measurements were used to extract 2 features. In combination, these extracted features were employed in fish image recognition of fish. For the purpose, hybrid meta-heuristic algorithms (Genetic Algorithm with Simulated Annealing Algorithm) together with back propagation classifier (MA-B Classifier) were applied in the classification of fish images into dangerous and non-dangerous families of fish, Images of dangerous fish are further recognized as Predatory or Poison fish family, whereas families of non-dangerous fish are classified into garden and food family. The feature extraction methods and classification algorithm considerably upgraded the recognition accuracy of BP as they improve and optimize the weights to be applied in its process of training and testing. The proposed BP and MA-B Classifiers show superior performance over the state of the art methods with accuracy percentage of 82.25 and 90 respectively.

REFERENCES

[1] A. M. Al Smadi, M. K. Alsmadi, H. Al Bazar, S. Alrashed, and B. S. Al Smadi, "Accessing Social Network Sites Using Work Smartphone for Face Recognition and Authentication," *Research Journal of Applied Sciences, Engineering and Technology,* vol. 11, pp. 56-62, 2015.
[2] M. Alsmadi, "Facial recognition under expression variations," *Int. Arab J. Inf. Technol.,* vol. 13, pp. 133-141, 2016.
[3] M. Alsmadi, K. Omar, and I. Almarashdeh, *Fish Classification: Fish Classification Using Memetic Algorithms with Back Propagation Classifier,* LAP Lambert Academic Publishing, 2012.
[4] M. K. Alsmadi, "An efficient similarity measure for content based image retrieval using memetic algorithm," *Egyptian Journal of Basic and Applied Sciences,* vol. 4, no. 2, pp. 112-122, Jun 2017.
[5] M. K. Alsmadi, "Query-sensitive similarity measure for content-based image retrieval using meta-heuristic algorithm," *Journal of King Saud University-Computer and Information Sciences,* vol. 30, no. 3, pp. 373-81, 2018.
[6] M. K. Alsmadi, *Facial expression recognition,* ed: Google Patents, 2018.
[7] M. k. Alsmadi, K. B. Omar, and S. A. Noah, "Proposed method to decide the appropriate feature set for fish classification tasks using Artificial Neural Network and Decision Tree," *IJCSNS,* vol. 9, pp. 297-301, 2009.
[8] T. H. Farag, W. A. Hassan, H. A. Ayad, A. S. AlBahussain, U. A. Badawi, and M. K. Alsmadi, "Extended Absolute Fuzzy Connectedness Segmentation Algorithm Utilizing Region and Boundary-Based Information," *Arabian Journal for Science and Engineering,* pp. 1-11, 2017.
[9] M. Alsmadi, K. Omar, S. Noah, and I. Almarashdeh, "A hybrid memetic algorithm with back-propagation classifier for fish classification based on robust features extraction from PLGF and shape measurements," *Information Technology Journal,* vol. 10, pp. 944-954, 2011.
[10] M. Alsmadi, K. Omar, S. Noah, I. Almarashdeh, S. Al-Omari, P. Sumari, S. Al-Taweel, A. Husain, N. Al-Milli, and M. Alsmadi, "Fish recognition based on robust features extraction from size and shape measurements using neural network," *Information Technology Journal,* vol. 10, pp. 427-434, 2009.
[11] M. Alsmadi, K. B. Omar, and S. A. Noah, "Back propagation algorithm: the best algorithm among the multi-layer perceptron algorithm," *International Journal of Computer Science and Network Security,* vol. 9, pp. 378-383, 2009.
[12] M. K. Alsmadi, "A hybrid firefly algorithm with fuzzy-C mean algorithm for MRI brain segmentation," *Am. J. Applied Sei,* vol. 11, pp. 1676-1691, 2014.
[13] M. K. Alsmadi, "MRI brain segmentation using a hybrid artificial bee colony algorithm with fuzzy-c mean algorithm," *Journal of Applied Sciences,* vol. 15, pp. 100, 2015.
[14] M. K. Alsmadi, "A hybrid Fuzzy C-Means and Neutrosophic for jaw lesions segmentation," *Ain Shams Engineering Journal,* 2017.
[15] M. K. Alsmadi, *Apparatus and method for lesions segmentation,* ed: US Patent App. 15/614,893, 2018.
[16] M. K. Alsmadi, K. B. Omar, and S. A. Noah, "Fish recognition based on robust features extraction from size and shape measurements using back-propagation classifier," *International Review on Computers and Software,* vol. 5, pp. 489-494, 2010.
[17] M. K. Alsmadi, K. B. Omar, and S. A. Noah, "Fish classification based on robust features extraction from color signature using back-propagation classifier," *Journal of Computer Science,* vol. 7, pp. 52, 2011.
[18] M. K. Alsmadi, K. B. Omar, S. A. Noah, and I. Almarashdeh, "Fish recognition based on robust features extraction from color texture measurements using back-propagation classifier," *Journal of Theoritical and Applied Information Technology,* vol. 18, 2010.
[19] M. K. S. Alsmadi, K. B. Omar, S. A. Noah, and I. Almarashdah, "Fish recognition based on the combination between robust feature selection, image segmentation and geometrical parameter techniques using Artificial Neural Network and Decision Tree," *arXiv preprint arXiv:0912.0986,* 2009.
[20] Z. Thalji and M. Alsmadi, "Iris Recognition using robust algorithm for eyelid, eyelash and shadow avoiding," *World Applied Sciences Journal,* vol. 25, pp. 858-865, 2013.

[21] T. T. Hnin and K. T. Lynn, "Fish Classification Based on Robust Features Selection Using Machine Learning Techniques," in *Genetic and Evolutionary Computing: Proceedings of the Ninth International Conference on Genetic and Evolutionary Computing, August 26-28, 2015, Yangon, Myanmar - Volume 1, T. T. Zin, J. C.-W. Lin, J.-S. Pan, P. Tin, and M. Yokota, Eds., ed Cham: Springer International Publishing*, 2016, pp. 237-245.
[22] İ. B, K. Y, R. A.N, and T. C, "Image analysis methods on fish recognition," in *2014 22nd Signal Processing and Communications Applications Conference (SIU)*, 2014, pp. 1411-1414.
[23] S. Ogunlana, O. Olabode, S. Oluwadare, and G. Iwasokun, "Fish Classification Using Support Vector Machine," *African Journal of Computing & ICT*, vol. 8, pp. 75-82, 2015.
[24] S. Daramola and O. Omololu, "Fish Classification Algorithm using Single Value Decomposition," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 5, pp. 1621-1626, 2016.
[25] D. Mushfieldt, M. Ghaziasgar, and J. Connan, "Fish identification system," in *Proceedings of South African Telecommunication Networks and Applications Conference (SATNAC 2012)*, 2012, pp. 231-236.
[26] M.-C. Chuang, J.-N. Hwang, and K. Williams, "A feature learning and object recognition framework for underwater fish images," *IEEE Transactions on Image Processing*, vol. 25, pp. 1862-1872, 2016.
[27] J. Matai, R. Kastner, G. Cutter Jr, and D. Demer, "Automated techniques for detection and recognition of fishes using computer vision algorithms," in *NOAA Technical Memorandum NMFS-F/SPO-121, Report of the National Marine Fisheries Service Automated Image Processing Workshop, Williams K., Rooper C., Harms J., Eds., Seattle, Washington (September 4–7 2010)*, 2010.
[28] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, pp. 137-154, 2004.
[29] C. Pornpanomchai, B. Lurstwut, P. Leerasakultham, and W. Kitiyanan, "Shape and texture-based fish image recognition system," *Kasetsart Journal-Natural Science*, vol. 47, pp. 624-634, 2013.
[30] M. M. M. Fouad, H. M. Zawbaa, N. El-Bendary, and A. E. Hassanien, "Automatic nile tilapia fish classification approach using machine learning techniques," in *Hybrid Intelligent Systems (HIS), 2013 13th International Conference on*, 2013, pp. 173-178.
[31] B. S. Al Smadi, "Applications of Meta-Heuristic Algorithm with Back Propagation Classifier for Handling Class of General Fish Models," *International Journal of Computer Science and Network Security (IJCSNS)*, vol. 16, pp. 38, 2016.
[32] Z. Hu, Y. Zhang, Y. Zhao, L. Cao, Y. Bai, and M. Huang, "Fish eye recognition based on weighted constraint AdaBoost and pupil diameter automatic measurement with improved Hough circle transform," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 33, pp. 226-232, 2017.
[33] A. Baloch, M. Ali, F. Gul, S. Basir, and I. Afzal, "Fish Image Segmentation Algorithm (FISA) for Improving the Performance of Image Retrieval System," *International Journal of Advanced Computer Science and Applications*, vol. 3, pp. 396-403, 2017.
[34] A. Ali-Gombe, E. Elyan, and C. Jayne, "Fish Classification in Context of Noisy Images," in *International Conference on Engineering Applications of Neural Networks*, 2017, pp. 216-226.
[35] M. N. Mokti and R. A. Salam, "Hybrid of Mean-shift and median-cut algorithm for fish segmentation," in *Electronic Design, 2008. ICED 2008. International Conference on*, 2008, pp. 1-5.
[36] M. Alsmadi, K. B. Omar, S.A. Noah, I. Almarashdeh, "Fish Recognition Based on Robust Features Extraction from Size and Shape Measurements Using Neural Network " *Journal of Computer Science*, vol. 6, pp. 1088-1094, 2010.
[37] U. A. Badawi and M. K. Alsmadi, "A General Fish Classification Methodology Using Meta-Heuristic Algorithm With Back Propagation Classifier," *Journal of Theoretical & Applied Information Technology*, vol. 66, pp. 803-812, 2014.
[38] G. I. Sayed, A. E. Hassanien, A. Gamal, and H. A. Ella, "An Automated Fish Species Identification System Based on Crow Search Algorithm," *Cham*, 2018, pp. 112-123.
[39] V. Tutors, "Precalculus : Find the Measure of an Angle Between Two Vectors," 2018, [online], Available *https://*www.varsitytutors.com/precalculus-help/find-the-measure-of-an-angle-between-two-vectors.
[40] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *science*, vol. 220, pp. 671-680, 1983.
[41] H. Pirim, E. Bayraktar, and B. Eksioglu, "Tabu Search: A Comparative Study," W. Jaziri, Ed., ed Mississippi State University, Industrial and System Engineering Department USA: InTech, Croatia, 2008, pp. 1-28.
[42] F. Xhafa and A. Abraham, "Metaheuristics for Scheduling in Distributed Computing Environments," *ed: Springer*, vol. 146, 2008.
[43] R. Oliveira, T. Candeias, L. Santos, H. Shahbazkia, "Gabor filters optimized by simple simulated annealing," 2000.
[44] D.-M. Tsaia, S.-K. Wua, and M.-C. Chenb, "Optimal Gabor filter design for texture segmentation using stochastic optimization," *Image and Vision Computing*, vol. 19, pp. 299-316, 2001.
[45] M. k. Alsmadi, K. B. Omar, S. A. Noah, and I. Almarashdah, "Performance Comparison of Multi-layer Perceptron (Back Propagation, Delta Rule and Perceptron) algorithms in Neural Networks," in *2009 IEEE International Advance Computing Conference*, 2009, pp. 296-299.
[46] C. Lee, P. Panda, G. Srinivasan, and K. Roy, "Training deep spiking convolutional neural networks with STDP-based unsupervised pre-training followed by supervised fine-tuning," *Frontiers in neuroscience*, vol. 12, 2018.
[47] M. K. Alsmadi, "Forecasting River Flow in the USA Using a Hybrid Metaheuristic Algorithm with Back-Propagation Algorithm," *Scientific Journal of King Faisal University (Basic and Applied Sciences)*, vol. 18, pp. 13-24, 2017.

[48] U. A. Badawi and M. K. S. Alsmadi, "A Hybrid Memetic Algorithm (Genetic Algorithm and Great Deluge Local Search) With Back-Propagation Classifier for Fish Recognition " *International Journal of Computer Science Issues,* vol. 10, pp. 348-356, 2013.
[49] P. Galinier and J.-K. Hao, "Hybrid evolutionary algorithms for graph coloring," *Journal of combinatorial optimization,* vol. 3, pp. 379-397, 1999.
[50] C. Blum, L. Gaspero, A. Roli, M. Sampels, and A. Schaerf, *Hybrid Metaheuristics: 6th International Workshop, HM 2009 Udine, Italy, October 16-17, 2009 Proceedings*: *Springer*, vol. 5818, 2009.
[51] P. Moscato and C. Cotta, "A gentle introduction to memetic algorithms," in *Handbook of metaheuristics, ed: Springer,* 2003, pp. 105-144.
[52] K. C. Tan, T. H. Lee, D. Khoo, and E. F. Khor, "A multiobjective evolutionary algorithm toolbox for computer-aided multiobjective optimization," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* vol. 31, pp. 537-556, 2001.
[53] M. S. Nery, A. M. Machado, M. F. M. Campos, F. L. C. Padua, R. Carceroni, and J. P. Queiroz-Neto, "Determining the Appropriate Feature Set for Fish Classification Tasks," in *Computer Graphics and Image Processing, 2005. SIBGRAPI 2005. 18th Brazilian Symposium on,* 2005, pp. 173-180.
[54] C. Xie and Y. He, "Spectrum and image texture features analysis for early blight disease detection on eggplant leaves," *Sensors,* vol. 16, pp. 676, 2016.
[55] A. R. A. Majid, *Field Guide to Selected Commercial Marine Fishes of Malaysian Waters*: Fisheries Research Institute Department of Fisheries Malaysia, 2004.
[56] D. J. Lee, S. Redd, R. Schoenberger, X. Xiaoqian, and Z. Pengcheng, "An automated fish species classification and migration monitoring system," in *Industrial Electronics Society, 2003. IECON '03. The 29th Annual Conference of the IEEE,* vol. 2, pp. 1080-1085, 2003.
[57] D.-J. Lee, J. K. Archibald, R. B. Schoenberger, A. W, Dennis, and D. K. Shiozawa3, "Contour Matching for Fish Species Recognition and Migration Monitoring," in *Applications of Computational Intelligence in Biology*. *ed: springerlink*, pp. 183-207, vol. 122, 2008.