

Classification of Retina Diseases from OCT using Genetic Programming

Hadeel Abdulrahman
Postgraduate Student (MSc)
Department of Artificial Intelligence
Faculty of Informatics Engineering, Aleppo
University, Syria

Mohamed Khatib, PhD
Department of Artificial Intelligence
Faculty of Informatics Engineering, Aleppo
University, Syria

ABSTRACT

In this paper, a fully automated method for feature extraction and classification of retina diseases is implemented. The main idea is to find a method that can extract the important features from the Optical Coherence Tomography (OCT) image, and acquire a higher classification accuracy. The using of genetic programming (GP) can achieve that aim. Genetic programming is a good way to choose the best combination of feature extraction methods from a set of feature extraction methods and determine the proper parameters for each one of the selected extraction methods. 800 OCT images are used in the proposed method, of the most three popular retinal diseases: Choroidal neovascularization (CNV), Diabetic Macular Edema (DME) and Drusen, beside the normal OCT images. While the set of the feature extraction methods that is used in this paper contains: Gabor filter, Local Binary Pattern (LBP), Gray Level Co-occurrence Matrix (GLCM), histogram of the image, and Speed Up Robust Filter (SURF). These methods are used for the both of global and local feature extraction. After that the classification process is achieved by the Support Vector Machine (SVM). The proposed method performed high accuracy as compared with the traditional methods.

Keywords

Genetic programming, feature extraction, Optical Coherence Tomography, OCT image classification, OCT feature extraction.

1. INTRODUCTION

Several diseases may affect the retina, which is the main part of the eye and responsible for the vision process. The risk of retina diseases (such as CNV, DME and Drusen), ranges from visual impairment to vision loss, if not detected and treated. (see Figure 1) shows normal retina OCT and various retina diseases.

A useful medical imaging technique used recently in ophthalmology called OCT, makes the detect of these diseases easier. Because of its high-quality, it can be used to visualize retina layers accurately, so any abnormal change in its cells can be detected easily [1].

Genetic Programming is an evolutionary algorithm and also an extension of genetic algorithm, GP is able to produce human readable computer programs. It works by searching the space of possible computer programs for an individual computer program that is highly fit in solving or approximately solving the problem at hand [2].

The need of automated OCT diagnosis system is raised, because of the possibility of human diagnostics error, so simply the automated system can identify any normal image

from abnormal, and classify the abnormal image to one of the three diseases, by using a proper feature vector that is extracted using the proposed method.

2. RELATED WORKS

This subsection briefly surveys typical related works, which use GP and its related versions for feature extraction and classification, either for medical applications or for another applications.

For features extraction, some studies depend on GP to extract local features, global features, or both. By trying to find the best extraction methods which will be suitable for the studied application, or the best feature vector to reach the highest classification accuracy.

In [3], a GP used to design a new program structure for both local and global feature extraction, and synthesize these features as feature vector with flexible length that can be achieved a good classification performance.

A multi-layer GP approach for performing automatic region detection and feature extraction [4], using a set of image descriptor as function set. The main step is to find a set of regions of interest in the image, to do local feature detection, and to extract features from them, then a binary classification is done by subtracting the standard deviation of the selected regions.

In [5], the local features extracted by using LBP and SIFT of OCT images, then the feature vector is classified by random forest into normal and DME abnormal image.

The GP-HoG [6] consists of a constant number of GP tree layers, the bottom layer includes the HoG and terminal nodes, where terminals are the image and the coordinates for the region of interest. The middle layer consists of a mix of the

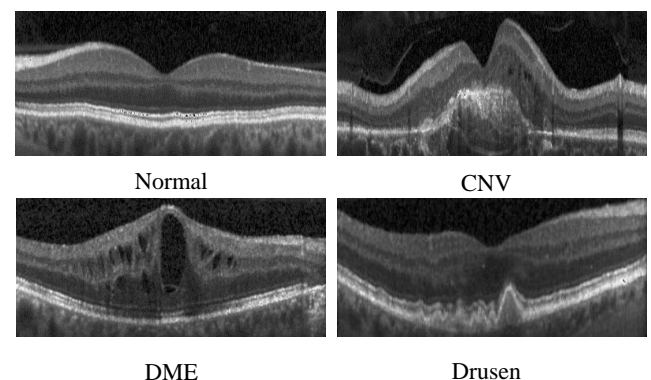


Fig 1: Normal OCT image and different diseases

bin and distance nodes. while, the top layer represents the classification part by using simple arithmetic operations.

Feature extraction step is accomplished by LBP for the skin cancer images dataset, then two GP stages are used, the first stage, for feature selection using LBP method, while the second stage for feature construction [7].

In [8], A new classification method based on combine SVM classifier with random forest to classify OCT images, using LBP feature extraction method.

3. THE PROPOSED METHOD

This section describes the overall methods, to find the proper program structure for feature extraction and OCT classification, including program representation, genetic operations, terminals and fitness function.

3.1 GP Program Representation

Strongly typed Genetic programming (STGP) is an enhanced version of genetic programming which enforces data type constraints and whose use of generic functions and generic data types makes it more powerful than other approaches to type constraint enforcement [9]. The proposed method depends on the STGP for the program representation, that contains the best combination of feature extraction methods and their parameters.

Feature extraction is one of the most important steps in computer vision, that enterprise helpful and brief information about the image, it can be used for classification or recognition issues. Features can be extracted locally or globally. Global features give information about the whole image, while local features give information from the important patches of the image. In the proposed method the both, local and global feature methods are used as functions for the GP tree nodes.

In [3], the root of the tree was the feature vector, while in the proposed method the root will produce a list of the selected extraction methods and their parameters. Terminals are the parameters of each extraction methods, and the coordinates of a patch for cropping the image to do the local feature extraction. (See Figure 2).

The length of the GP program is variable and depends on the count of the selected methods that will vary from two to six methods, and their parameters.

3.2 Function Set and Terminals:

A terminal usually contains variable inputs and constants, while a function set contains the program functions. A set of feature extraction methods are used as function in the GP program. these functions are: image histogram, Gabor filter, GLCM, LBP, and SURF. These methods are applied for both global and local feature extraction.

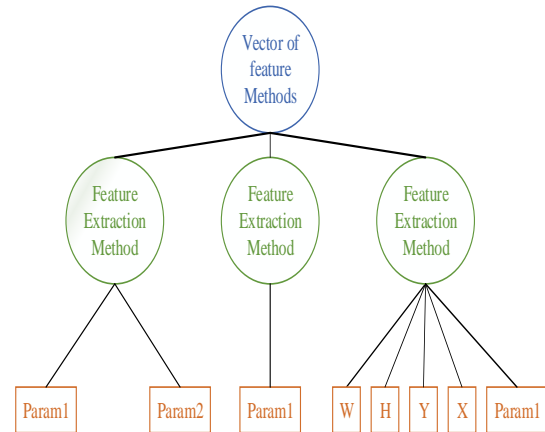


Fig 2: The proposed program structure representation

For global feature extraction, the extraction will be applied to the whole image, and the terminals of the GP-tree is each function parameters. While for the local feature extraction, the input of the function node is the coordinates of the top left point, height and width of the patch for cropping image, in addition to each function parameters. The local extraction will be applied for the cropped part of the image.

Histogram is a function that find the number of pixels for each intensity level in the image [10]. The parameter of the histogram that used in the proposed method is the number of bins, which is used to control the number of intensity levels and varied between 1 to 256, each bin represents the number of pixels in the image that have that intensity value. The number of the bins is selected randomly through the creation of the initial population. (see Table 1. for the extraction methods parameters).

Table 1. Feature extraction methods parameters

Feature Extraction Method	Parameters	Values
Histogram	Bin	[1,256]
Gabor filter	Scale	[2,5]
	Orientation	[4,8]
	Filter size	[32,39]
LBP	Radius	[1,5]
	Number of neighbors	[4,24]
SURF	Number of the strongest point	[1,10]
GLCM	Direction	0

Gabor filter [11] is a bank of filters used for texture analysis and also to extract important features from the image. The bank of filters consists of many filters, their count related to the parameters of the Gabor filter, like scale and orientation. Each filter has an orientation and scale differ from other filters. The value of the scale ranges between 2 to 5, while the

value of the orientation varies from 4 to 8 directions. The number of scales and orientations are selected randomly through the creation of the initial population. Also, the size of filters is chosen randomly, and varying between 32 to 39 since the details of the image are important.

LBP is a common way used to extract texture features from image. LBP scans image pixel by pixel using sliding window [12], with a variable radius that is generated randomly between 1 to 5 through the creation of initial population. The central pixel of the window is compared with its neighbors, and also the number of neighbors is a random selected parameter and may take values between 4 to 24. So, the result is a chain of binary numbers represent a gray level. Then the chain is converted to the histogram of these computed gray level, and this histogram is the required feature vector.

SURF is an image descriptor and detector, scale and orientation invariant, used to detect points of interest [13] in the overall image for global extraction method, or in some patches for local method, it depends on the integral image that speed up the calculation process. These points are defined by their orientation, location and scale, the number of points is varying from image to image, so in the proposed method a random number is selected of the strongest points of interest and that number varies between 1 to 10 points of interest. The feature vector from SURF is the information about orientation, location and scale from the selected points.

GLCM [14] also is a texture feature extraction method, the result of GLCM is a N square matrix of the image, where N is number of the image gray level. Each cell with (i,j) index represent the count of accompany between (i) intensity and (j) intensity, in the proposed method the direction 0 is used for GLCM calculation. From this matrix, some useful statistical values can be extracted and used as feature vector, like: Contrast, Correlation, Homogeneity and Energy.

Terminals of the GP tree, will be these functions' parameters, and the four coordinates of the image patch for local extraction.

3.3 Initial Population

The initial population is generated totally random. For each individual in the population, first of all, a number of methods is chosen, then for each one, a method is selected randomly from the methods list, and determined to be a global or local, after that, the parameters for every method are generated also randomly.

For the local method, in addition to its parameters, a four-numbers are generated (X, Y, H, W); X and Y are the coordinates of the top left cropping point, while H and W are the height and width of the cropping rectangle, with respect to the size of the input image.

3.4 Fitness Function

Each chromosome in the population is a program, the input for this program is the required parameters, while the output is a vector of selected feature extraction method. This vector is used to extract features from all the images in the dataset. These feature vectors are then feed as input to the SVM classifier, to classify the image to one of the four classes.

The dataset is divided into training set and testing set. So simply the fitness function of the program is the accuracy of the classifier. As shown in equation 1

$$Accuracy = \frac{TP+TN}{Total} * 100\% \quad (1)$$

Where, TP: true positive, the total number of positive images have been classified as positive; TN: true negative, the total number of negative images have been classified correctly; Total is the total number of images in the dataset.

3.5 Genetic Operations

Genetic programming evolves a population through genetic operations: crossover and mutation [2].

3.5.1 Cross Over Operation

Crossover is the main reproduce operation in genetic programming. Crossover generates new solutions from parent solution, that selected from the previous generation. The selection is done by the roulette wheel method, after the selection of two parent is done, a random number is generated between 0 and 1, and compared with the crossover probability, if the condition is fulfilled then the new solutions will be moved to the next generation, else the parents solution will be moved to the next generation.

Crossover is done by using the one-point cut method. A one random cut point is chosen randomly for each parent, to exchange information between the two parents and have two new children solutions.

3.5.2 Mutation Operation

Mutation is the process of making a little change in an existing solution. For each solution in the new generation, a random number will be generated between 0 and 1, and compared with the mutation probability, if the condition is fulfilled, then this solution will be mutated.

The mutation process is done by changing one of the function's parameters of the selected solution. So, a new solution will be reached.

3.6 Elitism

Elitism is a good way for transferring the best solution in the generation, to the next generation [2].

By this, the program performance will continue to improve or at least maintain its performance, since the best solution from the previous generation is moved to the next generation.

4. EXPERIMENT DESIGN

4.1 Dataset

The OCT dataset represents the most common medical retina patients presenting and receiving treatment at clinics [15].

The dataset from the study by Kermany et al, consists of 207.130 OCT images from 633 patients, belong to four classes: CNV, DME, Drusen, and normal.

During to the time that required to execute the proposed method, 200 OCT images were selected randomly from each OCT class. So, the total number of the used dataset is 800 images, with four different classes. Also, the image of the dataset is resized, to be 224*224 for each image.

4.2 Training and Testing Sets

To evaluate the proposed method, the k-fold cross-validation technique was used.

As mentioned before, A randomly 200 images were selected from each class of the dataset, and randomly split into 5 folds, to have a training and testing set, that contain the four classes.

4.3 GP Parameters

The GP parameters that used in this paper are mentioned in Table 2.

Table 2. Genetic programming parameters

Parameter	Value	Parameter	Value
Population size	100	Generation	100
Crossover rate	0.8	Selection	Roulette Wheel
Mutation rate	0.1	Elitism rate	0.1

The population size is set to be 100 individuals, and the initial population is generated randomly. Roulette wheel is applied for picking individuals from the previous generation to candidate the process of producing new generation, through a one-point cut crossover method, with 0.8 crossover rate. While the mutation rate is 0.1 and elitism rate is 0.1.

the process of producing new generation is repeated 100 times to have 100 generations or till to reach the fitness value of 99.9%.

The experiment on the dataset is repeated 10 times independently, and the average of the best fitness value from all repetition is calculated.

5. RESULTS AND DISCUSSIONS

This section compares the performance of the proposed method to the non-GP methods.

5.1 Classification Without GP Based Feature Extraction Methods

For each one of the feature extraction methods that used in the function set of the GP-tree, an independent experiment was done to extract features from the dataset with the 4 classes, by using the common parameters for each extraction method. The classification process is done by SVM to find the accuracy of this extraction method. This method is repeated 10 times to calculate the mean value of the accuracy.

The accuracy of the five experiments is compared between each other and between the proposed method, as showed in Table 3.

The four statistical values (Contrast, Correlation, Homogeneity and Energy) are extracted from GLCM, from the proposed dataset to form the feature vector, a classification is done by the SVM. And the mean accuracy for 10 runs is 46%.

Using 256-bins, a histogram of the dataset images is calculated, and used as the feature vector. The accuracy for 10 runs is 49%.

The locations, scales and orientations of the top 10 strongest point detected by SURF are extracted from the dataset, and used as feature vector, then classified by SVM, and the mean accuracy of this process is 59%.

LBP take 8 neighbors and the radius is 1 to find the feature vector, and the mean accuracy is 64%.

For Gabor filter, 40 filters are created with 5 scales and 8 different orientations, while the size of the filter is 39*39, the accuracy of this feature extraction method is 83%.

Through the traditional extraction methods, Gabor filter as global extractor, outperforms the other traditional methods significantly on the proposed dataset, and the accuracy will be improved, if the Gabor filter methods combined with other methods for both local and global feature extraction.

5.2 Classification Using GP Based Feature Extraction Methods

By using the GP parameters that mentioned before, the experiment is repeated 10 times independently, and the values of the best accuracy for each repetition is shown in the Table 4.

The mean value of the accuracy for 10 iterations, which represent the accuracy of the proposed method is 90.65%. And (Figure 3) shows the relation between the best fitness value of each run and the number of repetition (run).

Table 3. Feature extraction method comparison

Feature Extraction Method	Accuracy (%)
GLCM	46
Histogram	49
SURF	59
LBP	64
Gabor	83
The Proposed Method	90.65

Table 4. The accuracy of 10 repetition

Run	Accuracy (%)	Run	Accuracy (%)
1	91.125	6	90.25
2	90.25	7	90.50
3	91.25	8	91.125
4	90.75	9	90.50
5	90.25	10	90.50

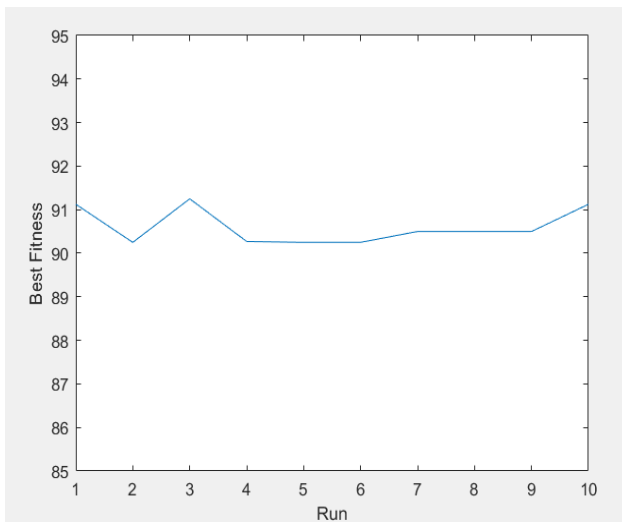


Fig 3: The values of the best fitness through 10 runs

According to the accuracy values of the 10 repetition, the accuracy of the proposed method is stable, and varies between 90.25% and 91.25%. This variation comes from the different solutions that acquired through the repetition, but these different solutions are similar to each other by the value of the accuracy.

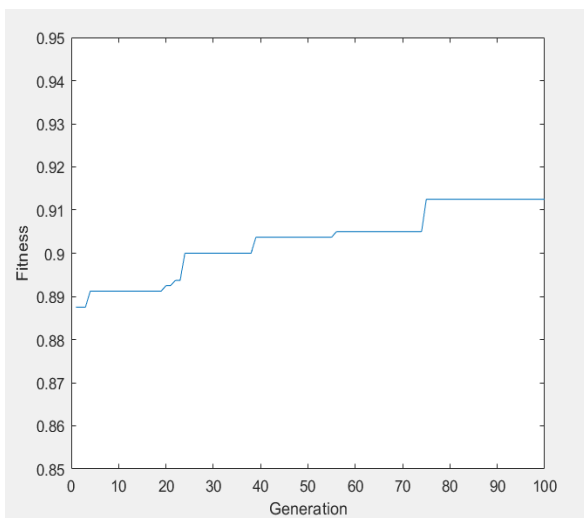


Fig 4: The values of fitness through 100 generation

(See Figure 4) shows the values of the fitness through 100 generations, and these values are either increasing or stable due to the elitism process.

Due to the random combination of feature extraction methods with random parameters value, the fitness value of the initial population is equal to 88.7%, and increased through the generation to reach 91.25%

The best program solutions that acquire the highest fitness value through the 10 repetition, shows the best combination of feature extraction and their parameters looks like:

Gabor (5 ,8, 32)

LBP (101, 2, 201, 59, 10, 4)

GLCM (115, 95, 137, 140)

LBP (30, 105, 97, 194, 18, 3).

Gabor filter will be used as global feature extractor, with 5 scales, 8 orientations and the size of all filters of the Gabor filter bank are 32*32.

LBP is a local extractor for the patch of the image with the coordinates (101,2,201,59), for 10 neighbors, and 4-radius.

GLCM is applied as local extractor for the patch of the image with the coordinates (115,95,137,140).

And LBP will be applied again for the patch (30,105,97,194), with 18 neighbors, and 3-radius.

(See Figure 5), shows the best GP tree solution. And (See Figure 6), that shows a DME image with three rectangles represent the regions of interest to extract a local features from them. The red rectangle represent the regoin of interset that important features is extracted from it using LBP. The yellow rectangle is the best regoin that can GLCM extract good features from it, while the green rectangle, is also another region that can LBP extract features from it.

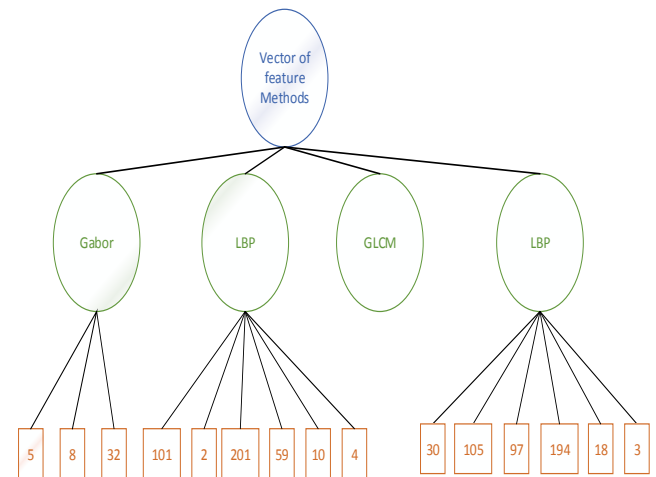


Fig 5: The best GP tree

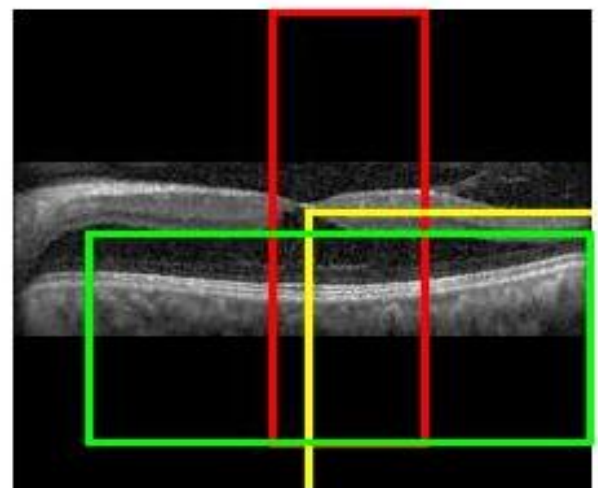


Fig 6: The regions of interest selected by the GP program

Features were extracted using a written code in MATLAB 2018a, on a personal computer (Microsoft Windows 10, Intel Core i7, CPU 1.80 GHz, 8.00 GB of RAM). While the classifier is the MATLAB SVM classifier with linear kernel function.

6. CONCLUSIONS

In this paper, an automated system that can classify the retina diseases was designed. Images that used were OCT images which is informative type of images. The proposed method extract feature from images through genetic programming, so it can find the best combination of global and local feature extraction methods.

The extraction was done for the four classes, and then classified by the SVM classifier, so the method can be applied for binary classification and multiclassification.

The proposed method achieved 90.65% of accuracy, and also compared with traditional methods, it showed good results compared to other methods.

7. REFERENCES

- [1] Adhi, M. and Duker, J. S., 2013. Optical coherence tomography – current and future applications. Wolters Kluwer Health, Lippincott Williams & Wilkins.
- [2] Koza, J. R., 1992. Genetic Programming on the Programming of Computers by Means of Natural Selection. The MIT Press Cambridge, Massachusetts London, England.
- [3] Bi, Y., Zhang, M., and Bing, X. 2018. Genetic Programming for Automatic Global and Local Feature Extraction to Image Classification. IEEE Congress on Evolutionary Computation.
- [4] Bi, Y., Zhang, M., and Bing, X. 2018. An Automatic Feature Extraction Approach to Image Classification Using Genetic Programming. Springer International Publishing AG, part of Springer Nature.
- [5] Dash, P., Sigappi, AN. 2018. Detection and Recognition of Diabetic Macular Edema from OCT Images Based on Local Feature Descriptor. International Journal of Pure and Applied Mathematics.
- [6] Lensen, A., Al-Sahaf, H., Zhang, M., and Xue, B. 2016. Genetic Programming for Region Detection, Feature Extraction, Feature Construction and Classification in Image Data. Springer International Publishing Switzerland.
- [7] UI Ain, Q. Xue, B., Al-Sahaf, H., and Zhang, M. 2018. Genetic Programming for Feature Selection and Feature Construction in Skin Cancer Image Classification. Springer Nature Switzerland AG.
- [8] Gholami, P., Hassani, M. S., Parthasarathy, M. K., Zelek, J., and Lakshminarayanan, V. 2018. Classification of Optical Coherence Tomography images for diagnosing different ocular diseases. Multimodal Biomedical Imaging XIII
- [9] Montana, D. 2002. Strongly Typed Genetic Programming.
- [10] Gonzalez, R., and Woods, R., Digital Image Processing, third edition, Prentice Hall, 2009.
- [11] Feichtinger, H., and Strohmer, T., 1998, Gabor Analysis and Algorithms: Theory and Applications, Birkhäuser.
- [12] Wang, L., and HE, DC., 1990, Texture Classification Using Texture Spectrum.
- [13] Bay, H., Tuytelaars, T., and Gool, L. V., 2008, SURF: Speeded Up Robust Features.
- [14] Haralick R.M., Shanmugan, K., and Dinstein, I., 1979. Statistical and Structural Approaches to Texture. Proc.IEEE 67:786-804.
- [15] Kermany, D., Zhang, K., and Goldbaum, M. 2018, Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification. Mendeley Data, v2