RECOD at ImageCLEF 2011: Medical Modality Classification using Genetic Programming

Fábio Augusto Faria, Rodrigo Tripodi Calumby, and Ricardo da Silva Torres

Institute of Computing - University of Campinas
Campinas, São Paulo, Brazil
{ffaria,tripodi,rtorres}@ic.unicamp.br
http://www.ic.unicamp.br

Abstract. This paper describes the participation of the RECOD group on the ImageCLEF 2011 Medical Modality Classification sub-task. We present an approach based on genetic programming and kNN for image classification. In our approach the genetic programming is used for the learning of good functions for the combination of similarities obtained from a set of global descriptors for different visual evidences such as color, texture, and shape. For each class of the dataset a combination function was learned and used as a kNN classifier. Final classification results were generated by a majority voting scheme with the voting functions from each class. Preliminary experiments have shown a good effectiveness of the approach and its potential for improvements.

Keywords: Genetic Programming, Medical Images, Image Classification, Pattern Recognition

1 Introduction

Currently with the advance of the technology of acquisition and storage of images along with the popularization of the Internet use we can notice the emerging of big image collections. Furthermore the increase of storage and processing power makes possible the use of intelligent computer systems for image manipulation based on machine learning. Machine learning techniques are used, for instance, on image processing, pattern analysis and recognition, and image retrieval, and classification [1].

In the medical field, the classification task is conducted for several types of images (e.g., x-ray, CT, and endoscopy) [2, 7, 4]. In [2], learning techniques are used for the detection and classification of benign and malign tumors and in [3] it is used to determine the gestational age of newborns. Moreover intelligent image classification approaches are used in several different fields such as diseases diagnosis [20], biometrics [6, 14], biology [23], and remote sensing [18].

In the image classification task, descriptors can be used to extract visual features from the images. A visual descriptor can be described by two functions; (1) an extraction algorithm that encodes visual characteristics (e.g., color, texture, and shape) into a feature vector; (2) a similarity measure (distance function) that computes the similarity of two images using their respective feature vectors. Hence different descriptors potentially describes different features for distinct visual evidences or even for the same evidence. For instance, the BIC [19] and JAC [24] descriptors characterize the color evidence by different means. One descriptor may be more effective for a dataset than another one, but no descriptor is perfect for all collections.

Therefore the use of learning techniques to combine different visual evidences has the objective of increasing the effectiveness of the classifiers by taking advantage of the power of different but potentially complementary descriptors. In this work we propose the use of Genetic Programming as a learning technique together with kNN classifiers (GP+kNN) for image classification tasks. This technique has been used in several application such as data mining, signal processing, image retrieval and classification [12, 5, 8, 26, 17].

The remaining of this text is organized as follows: Section 2 briefly describes the genetic programming technique. Section 3 presents our proposed approach for image classification using genetic programming. In Section 4 we present experimental setup and results for the ImageCLEF 2011 Medical Modality Classification. Finally, Section 5 presents our conclusions and future work.

2 Genetic Programming

Genetic Programming [13] is a machine learning technique based on biology evolutionary concepts, such as natural selection and survival. In this context each potential solution is seen as an individual evolving in a population. Evolutionary transformations are iteratively applied on the populations thus simulating sequential generations of individuals. Hence the individuals evolve by genetic transformations such as reproduction, mutation and crossover.

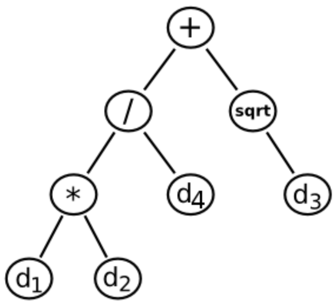
In general, GP individuals are encoded in a tree structure that represent programs that will evolve throughout the solution space towards a good solution for the problem. A fitness function is used to evaluate the individuals, and then it is possible to measure the quality of the solutions found in the evolutionary process. A basic evolution algorithm for genetic programming is presented in Algorithm 1.

Algorithm 1 Basic GP evolution algorithm.

```
1: Generate initial population of individuals
2: for N generations do
3:   Calculate the fitness of each individual
4:   Select the individuals to genetic operations
5:   Apply reproduction
6:   Apply crossover
7:   Apply mutation
8: end for
```

At first the initial population is generated, usually in random fashion (line 1). The iterative process starts for the evolution of the population through N generations (line 2). For each generation the fitness of the individuals are computed (line 3) and the genetic operators are then applied to the selected individuals. In order to evolve the population some individuals are properly selected (line 4) and they are subjected to genetic operators. The operators are responsible to introduce variability on the solutions making the moving through the solution space possible and consequently the discovery of better solutions. The reproduction operator just copies an individual to next generation (line 5). The crossover operator takes two individuals and exchange sub-trees from them creating two new individuals (line 6). The mutation operator just selects a random sub-tree from an individual and exchanges it for a new randomly generated sub-tree (line 7).

In our approach tree-based GP individuals are used to encode candidate functions that combine similarity measures obtained from different visual descriptors. In these trees inner nodes are arithmetic operators and the leaves are input nodes for similarity values computed using visual descriptors. Figure 1 presents an example combination function δD_{PG} that combines four different visual descriptors d_i.



(a)

$$\delta D_{PG} = \frac{d_1 * d_2}{d_4} + \sqrt{d_3}$$

(b)

Fig. 1. (a) Individual represented as a tree (b) Corresponding similarity function.

Details of our GP-based classification approach are presented in Section 3.

3 GP+kNN Classification Method

The approach for visual image classification of this paper was adapted from the one originally proposed in [25] for textual documents. Our approach is presented in Algorithm 2.

At the beginning the classifier is trained by the search of the best GP individuals (I_{c_i}), which means the search for the best similarity functions for each class c_i using the training set ($\mathcal{D}$) and the validation set ($\mathcal{V}$) (lines 1 to 9). For

Algorithm 2 *Approach for image classification using GP and kNN classifiers*

Require: Training set ($\mathcal{D}$), validation set ($\mathcal{V}$), and test set ($\mathcal{T}$), number of classes (N_c), number of generations (N_{gen})
1: **for all** $i \mid 1 < i \leq N_c$ **do**
2: Generate individuals populations (similarity functions) for each class c_i
3: **for all** $j \mid 0 < j \leq N_{gen}$ **do**
4: Calculate the fitness of the individuals using the fitness function and ($\mathcal{D}$)
5: Store N_{top} fittest individuals
6: Create a new population by: reproduction, crossover and mutation
7: **end for**
8: **end for**
9: Select the best I_{c_i} individual, among the $N_{gen} \times N_{top}$ candidates using ($\mathcal{V}$), which is unique for each class c_i
10: **for all** $w \mid 1 < w \leq \mathcal{T}$ **do**
11: **for all** $i \mid 1 < i \leq N_c$ **do**
12: $V_{w_i} \Leftarrow I_{c_i} + kNN$ applied to image w ($I_{c_i} + kNN$ is the kNN classifier using the individual of the class c_i)
13: **end for**
14: Class of $w \Leftarrow$ majority voting between all V_{w_i}
15: **end for**

this purpose $\mathcal{D}$ is used to find the $N_{gen} \times N_{top}$ candidates individuals and $\mathcal{V}$ is used to find the best individuals for each class. The classification step is then conducted on lines 10 to 15 for the test set ($\mathcal{T}$)

For the multi-class problem the algorithm uses N_c kNN classifiers. The final classification for each image w is decided by a majority voting using the output of all kNN classifiers.

3.1 Fitness Function

The fitness function used in our experiments was the FFP4 and its steps are presented in Algorithm 3. The bigger is the value of F the best is the similarity function.

Algorithm 3 FFP4 fitness function algorithm

F = 0
For each image D in class C
$Fitness_D = FFP4$ calculated using the $|C|$ images most similar to D
$F += Fitness_D$
$F = F/C$

The $FFP4$ equation is defined as:

$$FFP4 = \sum_{i=1}^{|C|} r(d_i) * k_8 * (k_9)^i \tag{1}$$

where $r(d)$ is the relevance of the image(1 if it is relevant, 0 otherwise). $|C|$ is the total number of images similar to D. For each image in class C, the fitness value is computed based on Eq. 1, and the final fitness is calculated as the mean FFP4 value for all images. The k_8=7.0 e k_9=0.982 were used according to the exhaustive analysis in [25].

4 Experiments

This section presents the methodology of the parametric search for the genetic programming approach and for the generation of final results for the ImageCLEF 2011 Medical Modality Classification Task.

4.1 Visual Descriptors

In this work we used a set of nine visual descriptors for global features such as color, texture and shape. Table 1 presents the descriptors used and the respective kind of evidence encoded.

Descriptor	**Visual Evidence**
ACC [10]	Color
BIC [19]	Color
CCV [16]	Color
CCOM [11]	Color
EOAC [15]	Shape
GCH [21]	Color
JAC [24]	Color
LAS [22]	Texture
QCCH [9]	Texture

Table 1. Visual image descriptors used in the experiments.

4.2 Data Setup

For the experiments on the genetic programming parametric search and for the final classification generation we created three different sub-datasets from the original train and test dataset of the classification task. On the parametric search the original training dataset from the task (989 images) was split into training (ps_train, 56%), validation (ps_val, 24%) and testing (ps_test, 20%) sub-datasets named *PSdataset*. With this three sub-datasets an exploratory search was conducted in order to find a good parametric setup for the genetic programming approach that should be used for original modality classification test set. The best configurations found are presented on Table 2. Table 2 presents the number of individuals on the population (Pop), the number of generations (# of Gen), the k factor on the kNN classifier, crossover, reproduction and mutation rates, and also the classification accuracy (Acc) obtained on the ps_test for each GP configuration.

Setup	Pop	# of Gen	kNN	Crossover	Reproduction	Mutation	Acc(%)
S1	10	10	1	0.9	0.1	0.7	68.78
S2	10	10	5	0.9	0.1	0.9	67.80
S3	10	10	7	0.1	0.2	0.7	68.78

Table 2. Best GP configurations found on the parametric search on the *PSdataset*.

For the final classification results for the original test set of the modality classification sub-task we used two different sub-datasets. The first one has the same training and validation sub-datasets of the *PSdataset* previously mentioned and was named as *Small*. The second one has the same validation sub-dataset of the *PSdataset* but at this time the training sub-dataset was composed by joining the ps_train and ps_test sub-datasets and was named *Large*. All sub-datasets were randomly generated from the original ones and Table 3 summarizes the amount of image samples for each sub-dataset used on the experiments.

Dataset	Training	Validation	Testing	Overall
PSdataset	547	237	205	989
Small	547	237	1024	1808
Large	752	237	1024	2013

Table 3. Number of images samples of the sub-datasets.

4.3 Final Results

For the ImageCLEF 2011 Medical Modality Classification sub-task our group submitted 10 runs with different sub-datasets and genetic programming configurations. The runs from 1 to 3 were generated using the *Small* sub-dataset with S1, S2 and S3 GP configurations, respectively. Similarly runs from 4 to 6 were generated using the *Large* sub-dataset.

For the runs 7 and 8 we conducted a majority voting on the results from 1 to 3 and 4 to 6, respectively. In case of ties the final classifications were decided by the classification present on runs 1 and 4, respectively.

Finally for runs 9 and 10 we conducted a majority voting on the results from 1 to 3 and 4 to 6, respectively. For this time, in case to ties the final classification was randomly chosen from one the voting runs.

Table 4 summarizes the configurations used for each run and their respective classification accuracy. The best run is highlighted.

Run	Setup	Run name	Accuracy(%)
1	S1	recod_imageclefmed_ModCla_343s	67.87
2	S2	recod_imageclefmed_ModCla_357s	66.69
3	S3	recod_imageclefmed_ModCla_370s	67.67
4	S1	recod_imageclefmed_ModCla_343l	67.48
5	**S2**	**recod_imageclefmed_ModCla_357l**	**69.72**
6	S3	recod_imageclefmed_ModCla_370l	67.87
7	S1,S2 and S3	recod_imageclefmed_ModCla_VsNoR	68.35
8	S1,S2 and S3	recod_imageclefmed_ModCla_VlNoR	69.04
9	S1,S2 and S3	recod_imageclefmed_ModCla_Vs	68.06
10	S1,S2 and S3	recod_imageclefmed_ModCla_Vl	69.43

Table 4. Classification results.

5 Conclusion

For the ImageCLEF 2011 Medical Modality Classification sub-task this work proposed the use of a learning technique for the combination of different visual evidences. The approach described in this paper is based on the use of genetic programming for the learning of effective similarity measures for the combination of visual similarities obtained from a set of global visual descriptors. For this purpose a kNN classifier using the similarity functions discovered was built for each class of the dataset and the final classification results were generated by a majority voting scheme.

In our experiments the GP+kNN classifiers were trained to combine visual evidences from images (e.g., color, texture, and shape). In the experiments the worst run submitted achieved an accuracy of 66.69% on the test set and our best run achieved 69.71%. Since we conducted only preliminary experiments we foresee a great potential of effectiveness growing of the approach, for example, by using more folds on the training step, using different visual descriptors and also incorporating textual information on the similarity functions.

Acknowledgments. FAPESP, CNPq and CAPES for financial support.

References

1. Andaló, F.A., Torres, R.d.S.: Descritores de forma baseados em tensor scale. Master's thesis, Instituto de Computação, Unicamp, Campinas, SP, Brazil (2007)
2. Antonie, M.L., Zaiane, O.R., Coman, A.: Application of Data Mining Tecniques for Medical Image Classification. In: 2nd Int. Workshop on Multimedia Data Mining MDM/KDD2001 (2001)
3. Araújo, A.V.e.a.: Aplicando Mineração de Imagem para Auxiliar na Determinação da Idade Gestacional em Recém-Nascidos. XI Congresso Brasileiro de Informática em Saúde, Florianópolis, Brazil (2006)
4. Bai, X., Qian, X.: Medical image classification based on fuzzy support vector machines. In: ICICTA '08: Proceedings of the 2008 International Conference on Intelligent Computation Technology and Automation. pp. 145–149. IEEE Computer Society, Washington, DC, USA (2008)
5. Bhanu, B., Lin, Y.: Object Detection in Multi-Modal Images Using Genetic Programming. Applied Soft Computing (2004)
6. Bowyer, K.W., Hollingsworth, K., Flynn, P.J.: Image understanding for iris biometrics: A survey. Computer Vision and Image Understanding 110(2), 281–307 (2008)
7. Cheng, P., Chien, B., Yang, W.: Medical image classification by supervised machine learning. In: AIC'06: Proceedings of the 6th WSEAS International Conference on Applied Informatics and Communications. pp. 106–110. World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA (2006)
8. Fan, W., Fox, E.A., Pathak, P., Wu, H.: The effects of fitness functions on genetic programming-based ranking discovery for web search: Research articles. J. Am. Soc. Inf. Sci. Technol. 55(7), 628–636 (2004)
9. Huang, C., Liu, Q.: An orientation independent texture descriptor for image retrieval. In: International Conference on Computational Science. pp. 772–776 (2007)
10. Huang, J., Kumar, R., Mitra, M., Zhu, W., Zabih, R.: Image indexing using color correlograms. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 762–768 (1997)
11. Kovalev, V., Volmer, S.: Color co-occurence descriptors for querying-by-example. In: Proceedings of the 1998 Conference on MultiMedia Modeling. pp. 32–38 (1998)
12. Koza, J.: Genetic Programming: On the programming of computers by natural selection. MIT Press (1992)
13. Koza, J.R.: Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press, Cambridge, MA, USA (1992)
14. Kumar, A., Passi, A.: Comparison and combination of iris matchers for reliable personal authentication. Pattern Recognition 43(3), 1016–1026 (2010)
15. Mahmoudi, F., Shanbehzadeh, J., Eftekhari-Moghadam, A., Soltanian-Zadeh, H.: Image retrieval based on shape similarity by edge orientation autocorrelogram. Pattern Recognition 36(8), 1725–1736 (2003)
16. Pass, G., Zabih, R., Miller, J.: Comparing images using color coherence vectors. In: ACM Multimedia. pp. 65–73 (1996)
17. da S. Torres, R., Falcão, A.X., Goncalves, M.A., Papa, J.P., Zhang, B., Fan, W., Fox, E.A.: A genetic programming framework for content-based image retrieval. Pattern Recognition 42(2), 283–292 (2009)
18. Santos, J.A., Torres, R.d.S.: Reconhecimento Semi-Automaático e Vetorização de Regiões em Imagens de Sensoriamento Remoto. Master's thesis, Campinas, SP, Brazil (2009)

19. Stehling, R., Nascimento, M., Falcão, A.: A compact and efficient image retrieval approach based on border/interior pixel classification. In: CIKM. pp. 102–109 (2002)
20. Suzuki, C.T.N.: Segmentação de Imagem de Parasitos Intestinais do Homem. Master's thesis, Campinas, SP, Brazil (2007)
21. Swain, M., Ballard, D.: Color indexing. International Journal of Computer Vision 7(1), 11–32 (1991)
22. Tao, B., Dickinson, B.: Texture recognition and image retrieval using gradient indexing. Journal of Visual Communication and Image Representation 11(3), 327–342 (2000)
23. Tofilski, A., Wing, D.: A program for numerical description of insect wings. Journal of Insect Science 4(17), 1–5 (2004)
24. Williams, A., Yoon, P.: Content-based image retrieval using joint correlograms. Multimedia Tools and Applications 34(2), 239–248 (2007)
25. Zhang, B., Chen, Y., Fan, W., Fox, E.A., Gonçalves, M., Cristo, M., Calado, P.: Intelligent gp fusion from multiple sources for text classification. In: CIKM '05: Proceedings of the 14th ACM international conference on Information and knowledge management. pp. 477–484. ACM, New York, NY, USA (2005)
26. Zhang, B., Gonçalves, M.A., Fan, W., Chen, Y., Fox, E.A., Calado, P., Cristo, M.: Combining structural and citation-based evidence for text classification. In: CIKM '04: Proceedings of the thirteenth ACM international conference on Information and knowledge management. pp. 162–163. ACM, New York, NY, USA (2004)