Genetic Programming with A New Representation and A New Mutation Operator for Image Classification

Qinglan Fan, Ying Bi, Bing Xue, and Mengjie Zhang
School of Engineering and Computer Science
Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand
{Qinglan.Fan;Ying.Bi;Bing.Xue;Mengjie.Zhang}@ecs.vuw.ac.nz

ABSTRACT

Due to the high dimensionality and variations of the image data, it is challenging to develop an image classification method that is able to capture useful information from images and then conduct classification effectively. This paper proposes a new GP approach to image classification, which can perform feature extraction, feature construction, and classification simultaneously. The new approach can extract and construct multiple informative features to effectively handle image variations. Furthermore, a new mutation operator is developed to dynamically adjust the size of the evolved GP programs. The experimental results show that the proposed approach achieves significantly better or similar performance than/to the baseline methods on two datasets.

KEYWORDS

Image Classification, Genetic Programming, Representation

ACM Reference Format:
Qinglan Fan, Ying Bi, Bing Xue, and Mengjie Zhang. 2021. Genetic Programming with A New Representation and A New Mutation Operator for Image Classification. In *2021 Genetic and Evolutionary Computation Conference Companion (GECCO '21 Companion), July 10–14, 2021, Lille, France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3449726.3459468

1 INTRODUCTION

Image classification is a promising area within computer vision. The task is to classify images into different groups according to the content in images. The main challenges to improve classification accuracy are the high dimensionality and variations of the image data, such as viewpoint, scale, and deformation variations.

Feature extraction, as one essential step of image classification, can effectively reduce image data dimensionality. However, most existing feature extraction methods are designed to conduct specific tasks and might not be effective for complex image classification tasks. Feature construction is capable of building multiple high-level features for capturing different parts of the image data, which can potentially improve the classification performance. However, developing effective feature extraction and construction methods for image classification is not a easy task due to image variations. Deep convolutional neural networks have gained much success in feature extraction and image classification, but these methods are computationally expensive, require a large amount of training data, and have poor interpretability.

One promising technique for image classification is genetic programming (GP), which uses evolutionary principles to automatically evolve/learn computer programs (solutions) to conduct image classification. It has been successfully applied to image classification and achieved promising results. However, the most existing methods either can address binary image classification tasks only, or depend on a predefined/preset algorithm such as support vector machine (SVM) to perform multi-class classification while using GP to perform feature extraction. This might limit their flexibility since the most effective combination of features and classifiers is unknown. Furthermore, one of the problems for GP is that the size of an evolved program can become very large without significant improvement in fitness.

In this paper, we will develop a new GP approach with a new program representation and a new mutation operator for binary and multi-class image classification tasks. The evolved GP programs can conduct image filtering, feature extraction, feature construction, and classification, automatically and simultaneously. Moreover, a new mutation operator will be proposed to evolve GP programs with suitable size. This approach will be examined on two datasets with different classification tasks (binary and multi-class) and compared with a number of baseline methods.

2 THE PROPOSED APPROACH

2.1 The New GP Program Representation

The developed GP program representation is based on strongly typed GP. It has multiple layers with different functions, which can perform image filtering, feature extraction, feature construction, and classification, automatically and simultaneously. The filtering layer of the GP program is flexible. Feature extraction aims to extract useful features from the raw image or filtered image. Feature construction is to produce new high-level features using extracted features. Classification can automatically select an effective classification algorithm such as SVM for assigning an image with a class label.

2.1.1 Function Set. The function set contains nine filtering functions, seven feature extraction methods, four feature construction methods, and one classification selection method. The nine filtering functions are $Median$, $Mean$, Min, Max, Gau, $GauD$, Lap, $LoG1$, and $LoG2$ that can effectively denoise, detect edges or flat region of the image. The seven feature extraction methods can capture different features from the image such as texture and edge. They are $uLBP$, $Gabor$, $GLCM$, HOG, $SIFT$, $Sobel$, and $Prewitt$. The four feature construction functions are $Feastures2W$ and $Features2C/3C/4C$.

GECCO '21 Companion, July 10–14, 2021, Lille, France

ACM ISBN 978-1-4503-8351-6/21/07.
https://doi.org/10.1145/3449726.3459468

$Feastures2W$ function can produce new high-level features by adding two weighted vectors into a vector. $Features2C/3C/4C$ is to concatenate two/three/four feature vectors into a feature vector. For the classification layer, the $classifiers$ can automatically select a classification algorithm from a set of classification algorithms including SVM, LR, RF, and ERF to generate the effective combinations of features and classification algorithms.

2.1.2 Terminal Set. Terminals are raw images or parameters of functions. The *Image* terminal represents the input image, which is a 2D array and the pixel values in the array are in the range [0, 1]. σ is standard deviation of the Gaussian filtering, which is in {1, 2, 3}. o_1, o_2 are the order of the Gaussian derivatives, which are in {0, 1, 2}. θ is the orientation of the Gabor filtering. Its value range is $[0, \frac{7\pi}{8}]$ with a step of $\frac{\pi}{8}$. $f = \frac{\frac{\pi}{2}}{\sqrt{2}^v}$ is the frequency of the Gabor filtering, where v is in {1, 2, 3, 4}. w is the weight coefficient of the feature vector.

2.2 The New Mutation Operator

Considering the size of individuals in conjunction with their corresponding fitness, the size or complexity of offsprings generated by the new mutation operator is expected to be consistent with the complexity of problems after conducting the mutation operation. For example, if the best individual expresses the large size, the size of the offsprings is expected to grow compared with their parents. However, if the best individual has a small size, the size of the offsprings is expected to be smaller than their parents.

3 EXPERIMENTS AND RESULTS

The performance of the proposed approach is examined on two image classification datasets of varying difficulty. FEI_1 [4] is a facial expression classification dataset, including facial images with smiling or natural expressions from different people. It has 150 training instances and 50 test instances. The number of instances per class is same. The FS [2] is to classify different natural scene images into 13 groups, including street, highway, coast, mountain, etc. It has 1300 training instances (100 instances per class) and 2559 test instances.

A number of baseline methods are used for comparisons, including traditional classification algorithms, GP-based methods and neural network-based methods. Four traditional classification algorithms are SVM, logistic regression (LR), random forest (RF), extremely randomized trees (ERF) that take the normalized raw image pixels as the input. FGP [1] is a recent GP-based image classification method. In this method, a feature vector can be produced by the evolved GP program, which is the input of SVM for classification. Neural network-based methods are LeNet-5 [3] and CNN-5 [1]. In the proposed approach, the maximum number of generation is 50, and the population size is set to 500. The crossover, mutation, and elitism rates are 0.8, 0.19, and 0.01, respectively. The tree depth at the first generation is between 2-15. Tournament selection with size 7 is used for selection. The experiments of all methods on each datasets independently run 30 times with different random seeds.

The maximum classification accuracy: Max, the average classification accuracy of the 30 runs and the standard deviation: Mean and Std on the test sets are reported. The Wilcoxon rank-sum test with a 95% significance interval is employed to compare the proposed approach with a benchmark method. The symbols "+" and "−" in Table 1 mean that the performance of the proposed approach is significantly better and worse than that of compared method. The symbol "=" indicates that the results of the proposed approach and the compared method are similar.

Table 1: Classification accuracy(%) of the proposed approach and compared methods on FEI_1 and FS

methods	FEI_1		FS	
	Mean±Std	Max	Mean±Std	Max
SVM	90.00 ± 0.00 +	90.00	20.30 ± 0.19 +	20.63
LR	92.00 ± 0.00 +	92.00	23.49 ± 0.00 +	23.49
RF	97.07 ± 1.00 -	98.00	36.53 ± 0.48 +	37.36
ERF	93.27 ± 0.96 +	94.00	37.15 ± 0.35 +	37.94
LeNet-5 [3]	95.27 ± 2.22 =	98.00	44.33 ± 1.66 +	48.53
CNN-5 [1]	94.13 ±1.36 =	96.00	48.01 ± 2.34 +	52.79
FGP [1]	93.60 ± 2.60 +	98.00	70.59 ± 1.74 +	74.48
Ours	95.07 ± 1.69	98.00	74.51± 5.45	78.70
overall	4+, 2=, 1-		7+	

According to Table 1, the proposed approach achieves better or similar performance than/to benchmark methods on the two datasets with different classification tasks. SVM, LR, RF, and ERF perform differently when dealing with different datasets. They achieve competitive performance on the FEI_1 dataset, while the classification results are very unsatisfactory on the FS dataset. For example, SVM reaches 90.00% mean accuracy on the FEI_1 dataset but only obtains 20.30% mean accuracy on the FS dataset. A similar pattern can be seen in LeNet-5 and CNN-5. GP based methods have a more flexible representation, which means that evolved GP programs can have different sizes when solving various problems. FGP achieves promising classification results. Compared with FGP, the proposed approach reaches better performance. Especially for the complicated classification task, namely, FS (scene classification with 13 classes), the proposed achieves a 3.92% increase in the mean accuracy over FGP. The experimental results show that the developed approach has the ability to learn informative and useful features for different image classification tasks.

4 CONCLUSION

This paper developed a new GP based image classification approach in which the most effective combinations of features and classifiers can be produced. The proposed approach achieved better performance than baseline methods on two image datasets. In the future, we will continue to investigate the program structure and functions in GP to improve the classification performance.

REFERENCES

[1] Y. Bi, B. Xue, and M. Zhang. 2021. Genetic Programming With Image-Related Operators and a Flexible Program Structure for Feature Learning in Image Classification. *IEEE Transactions on Evolutionary Computation* 25, 1 (2021), 87–101.

[2] L. Fei-Fei and P. Perona. 2005. A Bayesian hierarchical model for learning natural scene categories. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 2. 524–531.

[3] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 11 (1998), 2278–2324.

[4] C. E. Thomaz. 2012. FEI Face Database. https://fei.edu.br/~cet/facedatabase.html.