

Automatically Extracting Features for Face Classification Using Multi-Objective Genetic Programming

Ying Bi, Bing Xue, and Mengjie Zhang

School of Engineering and Computer Science

Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand

{Ying.Bi;Bing.Xue;Mengjie.Zhang}@ecs.vuw.ac.nz

ABSTRACT

Automatic feature extraction can be solved by maximising the classification accuracy and minimising the number of extracted features. This work proposes a new multi-objective feature extraction algorithm using genetic programming (GP) for face classification. The new multi-objective GP-based feature extraction algorithm with the idea of non-dominated sorting, which aims to maximise the objectives of the classification accuracy and the ratio of the number of extracted features. The results show that the proposed algorithm achieves significantly better performance than the baseline methods on two different face classification datasets.

KEYWORDS

Genetic Programming, Face Classification, Feature Extraction, Evolutionary Multi-Objective

ACM Reference Format:

Ying Bi, Bing Xue, and Mengjie Zhang. 2020. Automatically Extracting Features for Face Classification Using Multi-Objective Genetic Programming. In *Genetic and Evolutionary Computation Conference Companion (GECCO '20 Companion)*, July 8–12, 2020, Cancún, Mexico. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3377929.3389989>

1 INTRODUCTION

Face classification is a task of classifying the face of a person into one of predefined classes. The task is challenging due to the wide variations of pose, illumination, ageing, expression, and occlusion. Typically, image features are manually extracted, where the process requires human intervention and domain knowledge. Automatically extracting features from images is more effective for face classification. Such methods often achieve better classification performance than the methods using manually extracted features. However, the majority of the existing methods are based on neural networks (NNs), which often require a large number of training instances. Instead of using NNs, this study proposes a non-NN-based algorithm using genetic programming (GP) to achieve automatic feature extraction for face classification.

Many traditional face classification methods perform dimensionality reduction to obtain a small number of features for classification. A smaller number of features can not only shorten the training time

of a classification algorithm but also have higher interpretability. However, the NN-based methods tend to learn high-dimensional representation in order to achieve good classification performance. Typically, the objective of maximising the classification performance and the objective of minimising the number of extracted features are two conflicting objectives. The task of automatic feature extraction can be formulated as a multi-objective optimisation problem and solved using an evolutionary multi-objective algorithm. In this paper, we develop a multi-objective GP algorithm for automatic feature extraction in face classification.

The overall goal of this paper is to develop a new multi-objective feature extraction algorithm for face classification using genetic programming with the objectives of maximising the classification performance and minimising the number of extracted features. The multi-objective GP algorithm uses the idea of non-dominated sorting to search for a set of non-dominated solutions. This method will be examined on two face classification datasets and compared with the baseline methods.

2 THE PROPOSED APPROACH

2.1 Representation

The representation of the proposed approach is based on strongly typed GP. It integrates the processes of region selection, feature extraction and feature combination into a single tree. Region selection aims to select regions from the input image. Feature extraction is to extract features from the selected regions. Feature combination is to combine the extracted features to produce a combination of various features.

2.1.1 Terminal Set. The terminal set has the *Image*, *X*, *Y*, *S*, *W*, and *H* terminals. *Image* represents the input image, from where the features are extracted. $X \in [0, Image_width - 20]$ and $Y \in [0, Image_height - 20]$ represent the coordinates of the top-left point of the selected region in the image. *S*, *W* and *H* ($\in [20, 50]$) terminals represent the size of a region selected by the region selection function *RSS* or *RSR*.

2.1.2 Function Set. The functions are two region selection functions, four feature extraction functions and two feature combination functions. The two region selection functions are the *RSS* and *RSR* functions, which can select a square and rectangle region from the input face image, respectively. The four feature extraction functions are *LDA* [2], *SIFT* [5], *LBP* [1], and *Conca*, which are commonly used descriptors. The two feature combination functions are the *FC2* and *FC3* functions, which combine features extracted from 2 and 3 regions, respectively. These two functions can be the root node or the children nodes of a GP tree.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

GECCO '20 Companion, July 8–12, 2020, Cancún, Mexico

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-7127-8/20/07.

<https://doi.org/10.1145/3377929.3389989>

2.2 Multi-objective GP for Automatic Feature Extraction

2.2.1 Objective function 1: the classification accuracy. The first objective function is the classification accuracy. The classification accuracy is calculated using five-fold cross-validation and a linear support vector machine (SVM) on the training set. In this process, the images in the training set are transformed into a set of features using a GP tree. Then feature normalisation is performed to rescale the extracted features using the min-max normalisation method. The normalised dataset are feed into the linear SVM using five-fold cross-validation and the mean test accuracy of the five folds is obtained as the objective value.

2.2.2 Objective function 2: the number of extracted features. Instead of using the number of extracted features as an objective function directly, we use the inverse ratio of the number of extracted features to change this objective from a minimisation problem to be a maximisation problem. The second objective is the ratio of the minimal number of extracted features of the number of extracted features. The value of this objective is in the range of [0, 1]. Based on the design of the algorithm, it can be found that the minimal number of the extracted features is the minimum value of $C - 1$ and 59, where $C - 1$ indicates the number of features extracted by the LDA function and 59 indicates the number of features extracted by the LBP function.

MOGP uses the idea of non-dominated sorting to search for a set of Pareto solutions. The framework of the MOGP algorithm is the same as that in the non-dominated sorting genetic algorithm II (NSGA-II) [3].

3 EXPERIMENTS AND RESULTS

The performance of MOGP is examined on two well-known face classification datasets of varying difficulty. The ORL dataset [6] has 400 gray-scale face images, i.e., ten images per class. In the experiments, five images per class are used to form the training set and the remaining images are used to form the test set. The Extended Yale B dataset [4] has 2424 face images of 38 different subjects. In the experiments, 20 images per class are used to form the training set and the remaining images are used to form the test set. The images in this dataset are resized to 90×100 to reduce the computational cost.

The performance of MOGP is compared with five traditional methods (Eigenfaces+SVM, Fisherfaces+SVM, SIFT+SVM, LBP+SVM, and Original+SVM) and a CNN-based method (LeNet). The Eigenfaces+SVM, Fisherfaces+SVM, SIFT+SVM, and LBP+SVM methods use different types of features as inputs to feed into the linear SVM for classification. The Original+SVM uses raw pixel values for classification. In MOGP, the maximum number of generations is set to 50 and the population size is set to 100. The elitism, crossover and mutation rates are 0.01, 0.8 and 0.19, respectively. Tournament selection with size 7 is employed for selection. The minimal tree depth is set to 2 and the maximal tree depth is set to 8. The experiments of each method on each dataset runs independent 30 times and the accuracy of the test sets is reported.

The maximum classification accuracy, average accuracy and standard deviation of the MOGP algorithm and the baseline methods are listed in Tables 1. The “↑”, “=” and “↓” symbols indicate that

the proposed MOGP algorithm achieves significantly better, similar and worse results than/to the compared algorithm, respectively. For MOGP, a set of non-dominated solutions are obtained in each run. To compare it with the baseline methods, the classification results of MOGP are obtained using the individual with the best classification performance of the training set. The best results on each dataset are highlighted in bold.

Table 1: Test accuracy (%)

| Data Set | Max | Mean±St.dev | Max | Mean±St.dev |
|-----------------|--------------|-------------------|-----------------|-------------------|
| | ORL | | Extended Yale B | |
| Eigenfaces+SVM | 96.50 | 96.00±0.32+ | 77.04 | 76.02±0.54+ |
| Fisherfaces+SVM | 92.50 | 92.47±0.13+ | 80.47 | 77.93±2.08+ |
| SIFT+SVM | 97.00 | 97.00±0.00+ | 73.68 | 73.68±0.00+ |
| LBP+SVM | 91.00 | 89.23±0.75+ | 49.34 | 42.36±4.08+ |
| Original+SVM | 95.50 | 95.50±0.00+ | 90.81 | 90.75±0.03+ |
| LeNet | 93.50 | 88.33±2.91+ | 93.33 | 88.42±2.63+ |
| MOGP | 99.50 | 97.48±1.66 | 98.32 | 95.32±2.83 |

Table 1 shows that the proposed MOGP algorithm significantly outperforms any of the six baseline methods on the two face classification datasets. MOGP achieves not only the best maximum accuracy but also the best mean accuracy among all the methods on the two datasets. Specifically, MOGP improves the mean accuracy by 4.57% on the Extended Yale B dataset. The comparisons with the Eigenfaces+SVM, Fisherfaces+SVM, SIFT+SVM and LBP+SVM methods show that the features extracted by MOGP are more effective than the manually extracted features, i.e., the Eigenfaces, Fisherfaces, SIFT and LBP features. The comparisons with LeNet show that MOGP are more effective than LeNet for automatically extracting face features for classification. The experimental results demonstrate that MOGP is effective for face classification with simultaneously optimising the objectives of classification accuracy and the number of features.

4 CONCLUSIONS

This paper developed the MOGP algorithm to achieve automatic feature extraction for face classification with maximising the objectives of the classification accuracy and the inverse ratio of the number of extracted features. The results suggested that MOGP achieved significantly better classification performance than the baseline methods on the two datasets. In the future, we will investigate the performance of MOGP on other types of image classification tasks.

REFERENCES

- [1] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen. 2006. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 12 (2006), 2037–2041.
- [2] Peter N Belhumeur, João P Hespanha, and David J Kriegman. 1997. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 7 (1997), 711–720.
- [3] Kalyanmoy Deb, Samir Agrawal, Amrit Pratap, and Tanaka Meyarivan. 2000. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In *Proceedings of International conference on parallel problem solving from nature*. Springer, 849–858.
- [4] Kuang-Chih Lee, Jeffrey Ho, and David J Kriegman. 2005. Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5 (May 2005), 684–698.
- [5] David G Lowe. 2004. Distinctive image features from scale-invariant keypoints. *Proceedings of International Journal of Computer Vision* 60, 2 (2004), 91–110.
- [6] Ferdinando S Samaria and Andy C Harter. 1994. Parameterisation of a stochastic model for human face identification. In *Proceedings of the Second IEEE Workshop on Applications of Computer Vision*. 138–142.