Neurocomputing 172 (2016) 281–289

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

On random hyper-class random forest for visual classification

Teng Li [a,*], Bingbing Ni [b], Xinyu Wu [c], Qingwei Gao [a], Qianmu Li [d], Dong Sun [a]

[a] College of Electrical Engineering and Automation, Anhui University, Hefei 230039, China
[b] Advanced Digital Sciences Center, 1 Fusionoplis Way, # 08-10 Connexis North Tower, 138632, Singapore
[c] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
[d] School of Computer Science and Engineering, Nanjing University of Science and Technology, China

ARTICLE INFO

Article history:
Received 25 October 2013
Received in revised form
21 October 2014
Accepted 25 October 2014
Available online 9 May 2015

Keywords:
Random forest
Visual classification
Random hyper-class
Binary splitting
Multiple attributes

ABSTRACT

Random forest is an effective ensemble classifier. It has shown to be efficient with good generalization performance thanks to the randomness in training. However, for visual classification problems, where high dimensions and large number of classes exist, the impurity measurement in training the tree nodes split function not only neglects the strong conditional dependencies among visual attributes, but also leads to rather weak base classifiers, which may not reflect enough discriminative capability. Addressing this, we develop a random hyper-class random forest (RHC-RF) for visual classification tasks in this paper. During training each tree node, the entire class space is randomly partitioned into two hyper-classes and a weak learner targeting at separating the two hyper-classes is trained, where a linear classifier is learned with multiple feature attributes. This novel training scheme well preserves the randomness by random partition of the label space, and at the same time, the objective of training each node directly targets at discriminating classes in a *divide-and-conquer*-like manner. The proposed method is expected to benefit from the improved discrimination for visual classification compared with conventional random forests while preserving good generalization. Extensive experiments on various multi-class and high-dimensional visual classification tasks (including scene and object classification, image-based food identification, handwritten digit recognition and face recognition) demonstrate the superior accuracy, as well as its compactness and robustness achieved by the proposed method.

1. Introduction

In recent years, pattern classification methods with an ensemble of base learners have shown comparable and sometimes better than state-of-the-art methods in classification and regression. As one of the best ensemble classifiers, random forest has been successfully employed in several visual classification tasks [1–4], etc. It constructs a series of unpruned decision tree classifiers whose classification decisions are combined by a voting procedure in the end. Each base learner is randomized in two ways: first, each base-level decision tree classifier is learned with different randomly selected subsets drawn independently from the training set with replacement; and second, in each node of the tree the splitting attribute is selected from a randomly chosen sample of attributes.

A key step in constructing a randomized classification tree is training the branching function for each non-leaf node. For each non-leaf node, a split function is typically trained by employing a class-impurity measurement based on Gini index [5] (used in the well-known CART method) or entropy gain [6] (used in the well-known C4.5 algorithm). These impurity measurements are defined by how well the attribute splits the set into homogeneous subsets that have the same value of the target variable. These measurements reach minimum (zero) when all data samples in the node fall into a single target category. The algorithms that are used for constructing decision trees usually work top-down by choosing an attribute at each step, which is the next best attribute to use according the selected impurity measurement function, *e.g.*, Gini index or entropy gain, in splitting the set of items.

Offering good prediction performance, random forest methods are computationally effective and proven not to overfit, and are less sensitive to noisy data [7]. The effectiveness of random forest can be explained with the margin and correlation of diverse base classifiers learned in a bootstrapping style [8], and the prediction errors made by base classifiers should be uncorrelated so that the error correcting properties of the ensemble can have the maximum effect. The generalization capability of random forest can be attributed to the randomness injected during the training procedure. A theoretical study of the generalization error is given in [7]. An excellent review and tutorial of random forest methods is recently given in [9].

* Corresponding author.
E-mail addresses: tenglwy@gmail.com (T. Li), bingbing.ni@adsc.com.sg (B. Ni), xy.wu@siat.ac.cn (X. Wu), qingweigao@ahu.edu.cn (Q. Gao), liqianmu@126.com (Q. Li), sundong@ahu.edu.cn (D. Sun).

http://dx.doi.org/10.1016/j.neucom.2014.10.101
0925-2312/

Despite the success of random forests methodology, for visual classification tasks, where the feature dimensions are usually as high as thousands and the number of categories involved can be more than hundreds [1,10–13], conventional random forests cannot be guaranteed to reach the optimal performance. The impurity-based measures employed for nodes splitting in each base tree learning such as Gini index, though fast, assume the conditional independence of feature attributes [14], and evaluate each attribute separately while neglecting strong conditional dependencies among attributes. For high-dimensional visual categorization problems such dependencies usually arise thus the evaluation could result in poor performance of base classifiers. Since the base classifiers learned based on impurity measurement, they do not directly reflect the class discriminative capability, especially when the problem dimensionality is high (well-known as *curse of dimensionality*). Furthermore, it is possible for some tree nodes to have quite a limited number of samples for some classes, which impedes the class discrimination of base classifiers.

1.1. The proposed idea

In this work, inspired by the recent success of random label space partition method in metric learning for object recognition [15,16], we propose a Random Hyper-Class Random Forest (RHC-RF) for visual categorization tasks, aiming to improve the discriminative capability while keeping generalization of the random forest classifier. For splitting each non-leaf node in learning base tree classifiers, the proposed scheme *decomposes* the multi-class classification problem into a binary classification problem by random partition of the label space, and thus more readily to be solved by standard binary classifiers. Iteratively applying this scheme in the top-down growth of the decision tree can be considered as a *divide-and-conquer* scheme. Different from conventional random forests in tree node splitting, the randomness is not only injected from the random selection of attributes, but also the random partition of the label space. Moreover, several attributes can be selected and combined linearly as the split function, rather than simple impurity measures with a single attribute. The attributes are purely randomly selected, and they are linearly combined to construct the weak classifiers. Details will be given in Section 3.

This decision tree construction scheme can not only guarantee the discriminative partition of the data during each node splitting but also well preserve the randomness during the training which in turn guarantee the generalization capability of the constructed random forest. Theoretical analysis shows that random forest trained in this way leads to well-balanced random decision trees. Comprehensive experiments on different high-dimensional and multi-class visual classification problems including object and scene classification, face recognition, handwritten digit recognition, as well as image based fast food identification demonstrate the superiority of the proposed random forest in terms of accuracy, compactness and robustness.

The contributions of this paper are as follows: (1) an ensemble classifier based on popular random forest methods called RHC-RF is proposed, which handles the multi-class high-dimensional tasks and can be effective for visual classification; (2) the learned base decision trees are well-balanced, and the training and testing process of RHC-RF is efficient; (3) we demonstrate the proposed method on several visual classification tasks. It shows better performance than conventional random forests, while preserving good generalization performance.

The rest of this paper is organized as follows. Some related works are reviewed in Section 2. In Section 3, we present the details of our proposed Random Hyper-Class Random Forest method. The experimental results are given in Section 4 with comprehensive discussions. Finally, Section 5 concludes the paper.

2. Related works

The idea of ensemble learning, combining several different weak classifiers to have a single strong classifier in the end, has led to several popular classification methods, such as bagging [17], boosting [18], random forests [7] and error correcting output codes (ECOC) [19]. In the following we will give a brief review on ensemble learning for visual tasks, as well as the related random forest methods.

2.1. Ensemble learning for visual tasks

As the most popular algorithm for face detection, the AdaBoost, short for adaptive boosting algorithm [20], constructs a final classifier with weak classifiers learned serially through reweighting the training examples according to the accuracy, and in this way, the subsequent base learners gradually focus on the most problematic instances. The ensemble learning has also been applied for high-level features extraction for face verification [21]. Totally 65 binary classifiers are trained individually for visually describable attributes such as gender and race, as well as the similarities to a set of reference faces. The outputs of these classifiers are segregated as features for further verification.

As a closely relevant method first described by [22], ECOC segregates the multi-class learning problem to multiple binary problems to solve, for each a separate base classifier is trained. These problems are constructed by repeatedly partitioning the set of target classes into a pair of super-classes. Given a large number of such partitions, each target class can be represented as the codes representing the super-classes it belongs to. The classification is then performed by applying the base classifiers to generate the codes and make decision. Compared to the proposed random forest based method, ECOC has to choose a predicted class from a fixed set of classes.

These methods construct multiple base classifiers that rely on certain heuristic procedures or manual intervention, and often require large amount of training samples and high computation load. Bagging introduces randomness into the training algorithm. It makes bootstrap replicate sets which are randomly drawn from the original training set for base classifiers training, and the final decision is made by majority voting. Comparatively, random forests are ensembles of decision trees trained by bootstrapping but differ from bagging in that they also employ random feature selection.

A recent related work is random ensemble metrics (REMetric) proposed for metric learning for object recognition [15]. It applies the idea of random label space partition, and obtains multiple discriminative projection vectors from linear SVM using randomly sub-sampled training data. While achieving good performance, the REMetric, however, mainly focuses on metric learning without any consideration of random forest classifiers.

2.2. Related random forest methods

In visual classification tasks, random forests have been widely adopted and yielded comparable and sometimes better than state-of-the-art performance for different applications such as image classification [1], object detection [2], biomedicine [3], and 3-D human motion tracking [4]. Furthermore, random forest based methods also show the advantage of high computational efficiency over other popular classifiers such as support vector machine (SVM) [23]. However, the focus of these works is mainly on the effective feature representation, and they simply apply conventional random forests as the classifier without any modification.

A number of variants of random forest exist. Ho proposed several alternatives for the split functions [24]. One way is to compute the centroid of the c largest classes and that of the remaining points. A