Structure-Aware Face Clustering on a Large-Scale Graph with 10^7 Nodes

Shuai Shen[1,2], Wanhua Li[1,2], Zheng Zhu[1,2], Guan Huang[3], Dalong Du[3], Jiwen Lu[1,2,*], Jie Zhou[1,2]
[1] Department of Automation, Tsinghua University, China
[2] Beijing National Research Center for Information Science and Technology, China
[3] XForwardAI
{shens19,li-wh17}@mails.tsinghua.edu.cn;zhengzhu@tsinghua.edu.cn;
{guan.huang,dalong.du}@xforward.com;{lujiwen,jzhou}@tsinghua.edu.cn

Abstract

Face clustering is a promising method for annotating unlabeled face images. Recent supervised approaches have boosted the face clustering accuracy greatly, however their performance is still far from satisfactory. These methods can be roughly divided into global-based and local-based ones. Global-based methods suffer from the limitation of training data scale, while local-based ones are difficult to grasp the whole graph structure information and usually take a long time for inference. Previous approaches fail to tackle these two challenges simultaneously. To address the dilemma of large-scale training and efficient inference, we propose the STructure-AwaRe Face Clustering (STAR-FC) method. Specifically, we design a structure-preserved subgraph sampling strategy to explore the power of large-scale training data, which can increase the training data scale from 10^5 to 10^7. During inference, the STAR-FC performs efficient full-graph clustering with two steps: graph parsing and graph refinement. And the concept of node intimacy is introduced in the second step to mine the local structural information. The STAR-FC gets 91.97 pairwise F-score on partial MS1M within 310s which surpasses the state-of-the-arts. Furthermore, we are the first to train on very large-scale graph with 20M nodes, and achieve superior inference results on 12M testing data. Overall, as a simple and effective method, the proposed STAR-FC provides a strong baseline for large-scale face clustering. Code is available at https://sstzal.github.io/STAR-FC/.

1. Introduction

Recent years have witnessed the great progress of face recognition [9, 26, 27, 36, 37, 39]. Large-scale datasets are an important factor in the success of face recognition and there is an increasing demand for larger-scale data. Face clustering [22, 28, 40, 46, 48, 49, 50] is a natural way to solve the data annotation problem so as to make better use of massive unlabeled data. Face clustering is also one possible approach to organize and file large volumes of real face images in social media or other application scenarios.

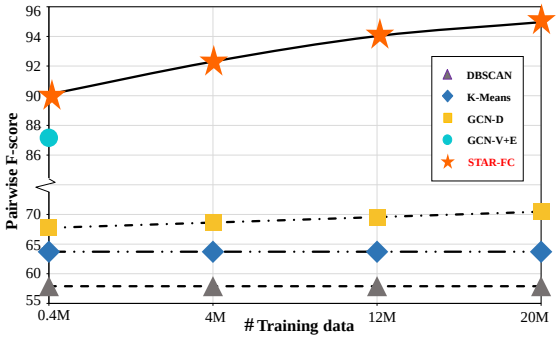

Figure 1: Method comparison when training with different scales of data and testing on 12M data from WebFace42M [53]. The proposed STAR-FC can fully explore the power of large-scale training data. GCN-V+E fails to handle larger training graph while GCN-D's performance is severely restricted due to the less consideration of the global structural information.

Recently a variety of efforts have been devoted to face clustering. Traditional unsupervised methods [16, 51] including K-Means [28] and DBSCAN [10] usually depend on some manually designed clustering strategies. They perform well on small datasets, however they are less effective when dealing with large-scale data as shown in Figure 1. Recent research trends [12, 40, 45, 47, 48] turn to the GCN-based supervised learning. These methods are performed based on the affinity graph and can be roughly divided into global-based and local-based ones *according to whether their GCN input is the whole graph or not*. The representative global-based method GCN-V+E [45] uses the entire graph for GCN training. As shown in Figure 1, it boosts the face clustering performance greatly compared with unsupervised methods, however the training data scale

* Corresponding author