PATENT SPECIFICATION
Confidential

TITLE OF THE INVENTION

Methods and apparatus for face recognition in the wild

INVENTORS

Shen Shengmei, Cheng Yu, Liu Qiankun
Pensees Pte Ltd
3 Fusionopolis Way, #14-23/24 Symbiosis, Singapore 138633

FIELD OF THE ART

The present invention relates to the field of face recognition and deep clustering using deep learning network. The method can provide a better recognition accuracy and boost the performance especially for large-scale face recognition applications.

The method can be used for face recognition applications where a gallery or black list dataset is given in prior. The method using clustering can give better representation for each identity among the dataset given; therefore, boost the performance of recognition.

BACKGROUND ART

The face recognition nowadays has achieved great boost in performance with deep learning technology compared to conventional methods using traditional machine learning where hand-crafted features are used together with signal and image analysis. The deep network is trained with big data and even able to recognize faces in extreme conditions.

However, with the increase of need for recognition under various unconstrained conditions especially when the required number of face ID to be covered becomes huge , the recognition task becomes more difficult. Unconstrained conditions include various pose, illumination, picture quality, with occlusion, wearing sunglasses or mask. The application for surveillance also requires the recognition among video clips.

Many researchers have been working on improving the performance of deep learning network in terms of structure and loss function. Although the performance is improved a bit, most of the methods directly use the features of its own test sample from learned model from deep learning network as the evaluation criteria, but lack of consideration of relationship between the testing samples.

SUMMARY OF THE INVENTION

Problems to be solved

The prior art as described above has the following problems.

Existing method only concentrate on improving the performance of deep learning framework but lack of exploring of the relationship among the samples in a database or among the test samples, which will can be reflected by the variance of the images of the same identity. The proposed method aims to excavate the relation among testing samples using clustering or other means to obtain a better representation thus achieve a better recognition accuracy, especially for face recognition of video clips.

Means of Solving the Problems

In the present invention, a general clustering framework is proposed as one mean for such relation in face recognition problems. The method contains two parts and suitable for various of recognition tasks.

1. The hard-clustering approach using single-link hierarchy clustering. With a properly chosen threshold, the samples are grouped into clusters and during evaluation, a voting scheme is applied to select the best match.
2. The soft-clustering approach using KNN (K-Nearest Neighbors) affinity matrix. With a properly chosen threshold, the sample features can be updated by pseudo-cluster center. The features of the same identity will move closer after the update.
3. The confidence re-balancing approach which utilizes the external information where face detection result is one way to lower down the similarity scores of low-confident samples to improve the true positive rate under very strict false positive rate.

Effects of Invention

The present invention provides a better face recognition performance using clustering as one mean. The testing features becomes more compact, where the features of the same identity shall have a smaller distance, which can improve the recognition accuracy at a large scale.

BEST MODE FOR CARRYING OUT THE INVENTION

With reference to the accompanying drawings, the preferred embodiments of the present invention are described as following.

FIG. 1 shows the general diagram for this invention, which is described here.

The face representation features are firstly extracted from the deep learning networks, then going through the said clustering module for post-processing. Users may choose one of the two approaches for different cases: hard-clustering for small dataset and soft-clustering for larger scales of database. After clustering process, the similarity score can be computed by the cosine similarity measurement or other means to calculate the similarity.
The confidence re-balance is another way to manipulate the said face features extracted from the previous step by utilizing other information, for example face detection result.

So the system can choose either clustering approach or confidence re-balance approach or choose the both approaches to achieve the best recognition or classification accuracy.

FIG. 2 shows the detail pipeline of hard-clustering approach. With given dataset in small size, we perform a hierarchy clustering with single-linkage with proper threshold to group the samples into several clusters. With a given probe sample, the similarity is computed by cosine similarity and a voting scheme is used. When the majority (>50% for example) of the cluster appears, i.e. the similarity score is above certain threshold, a label correction will be conducted for all the samples in the cluster.

FIG. 3 shows the details of soft-clustering approach. For large-scale data we suggest to use pseudo-center for each sample and do the clustering in an iterative manner.

FIG. 4 shows the details of confidence re-balancing. The low-confident samples are filtered out using other information, for example by the results of face detector where the confidence is below a user-defined threshold. When computing the pair-wise similarity scores, the scores are then reduced by a pre-defined value.

Industrial Applicability

The present invention can be applied to off-line face recognition processing, e.g. a large photo gallery or a large pool of face images. This method can be utilized as a plug-and-play module to efficiently enhance the performance of face recognition by post-processing.

BRIEF DESCRIPTION OF THE DRAWINGS

Tab.1: shows the outstanding performance of the present clustering method in Trillion Pairs dataset compared to the State-of-the-Art (SOTA) method for light recognition model (FLOPS<1G, model size<20M).
Tab.2: shows the outstanding performance of the present clustering method in Trillion Pairs dataset compared to the State-of-the-Art (SOTA) method for heavy recognition model (FLOPS<30G).
Tab.3: shows the outstanding performance of the present clustering method in IQiYi video dataset compared to the State-of-the-Art (SOTA) method for light recognition model (FLOPS<1G, model size<20M).
Tab.4: shows the outstanding performance of the present clustering method in IQiYi video dataset compared to the State-of-the-Art (SOTA) method for heavy recognition model (FLOPS<30G).
Tab.5: shows the results with confidence re-distribution and clustering on public IJB-C benchmark.

FIG. 1 shows a general diagram for the face feature clustering pipeline. Confidence re-balancing and clustering are two optional modules where users can choose to use either one or both.

FIG. 2 shows the detailed pipeline of hard-clustering. The testing samples are clustered using single-linkage clustering and finally a voting scheme is applied for each cluster. If the majority of samples appears (above certain threshold), all members in that cluster are assigned with the same identity.

FIG. 3 shows the details of the soft-clustering approach. The clustering is conducted in an iterative manner. In each step we find the affinity matrix and pseudo-center for updating the face features.

FIG. 4 shows the details of confidence re-balancing. The low-confident samples are selected by the face detector whose confident scores are below certain threshold. Then the similarity score is reduced by a pre-defined value if one of the samples is low-confident (connection CDE) and others keep the same (connection ABG).

WHAT IS CLAIMED IS:

1.Methods and apparatus for face recognition in the wild, comprising of:

Extracting features from input images and testing images in a test dataset, where features can be learned via deep learning model, can be hand-crafted features;

Examining the relationship among the testing images in a testing database by clustering the said testing images;

Adjusting and updating the image features according to the said cluster centers;

Evaluating the similarity between the said input image and each of the testing image in the said database by calculating the distance of their feature sets;

Determining if they belong to the same class or same ID for the case of face recognition.

2. Methods and apparatus for face recognition in the wild, comprising of, where Examining the relationship among the testing images in a testing database by clustering the said testing images in the claim 1, further comprising of:

Clustering the said test images by single-link hierarchy clustering, where the image features are grouped into clusters with a properly chosen threshold, and a voting scheme is applied to select the best match during the evaluation.

3. Methods and apparatus for face recognition in the wild, comprising of, where Examining the relationship among the testing images in a testing database by clustering the said testing images in the claim 1, further comprising of:

Clustering the said test images by using KNN (K-Nearest Neighbors) affinity matrix, where the image features can be updated by pseudo-cluster center with a properly chosen threshold, and the features belonging to the same class or identity will move closer after the updating.

4. Methods and apparatus for face recognition in the wild, comprising of, where Examining the relationship among the testing images in a testing database by clustering the said testing images in the claim 1, further comprising of:

The said clustering in the invention can be supervised and learned clustering model.

5.Methods and apparatus for face recognition in the wild, comprising of:

Extracting features from input images, where features can be learned via deep learning model, can be hand-crafted features;

Examining the confidence values of the input images and the testing images in a testing database by utilizing other information;

Calculating the similarity scores between the said input image and each of the testing images in the said database by calculating the distance of their feature sets;

Adjusting the said similarity score between the input image and each of the test images according to confidence re-balancing operation;

Determining if they belong to the same class or same ID for the case of face recognition.

6. Methods and apparatus for face recognition in the wild, comprising of, where Examining the confidence values of the input images and the testing images in a testing database by utilizing other information in the claim 5, further comprising of:

Evaluating the reliability of the input image and testing image by other tool including face detector, image quality assessment;

Determining the confidence values of the said input image and testing image, therefore the confidence values of the said image features of the input image and testing image, based on the above said operation.

7. Methods and apparatus for face recognition in the wild, comprising of, where Adjusting the said similarity score between the input image and each of the test images according to confidence re-balancing operation in the claim 5, further comprising of:

Reducing the said similarity score by a certain pre-defined value or learned value automatically, if the confidence value of one of the images is lower than a pre-defined threshold or learned value.

8.Methods and apparatus for face recognition in the wild, comprising of:

Extracting features from input images and testing images in a test dataset, where features can be learned via deep learning model, can be hand-crafted features;

Examining the relationship among the input images by clustering the said input images;

Adjusting and updating the image features according to the said cluster centers;

Evaluating the similarity between the said input image and each of the testing image in the said database by calculating the distance of their feature sets;

Determining if they belong to the same class or same ID for the case of face recognition.

9. Methods and apparatus for face recognition in the wild, comprising of, where Examining the relationship among the input images by clustering the said input images in the claim 8, further comprising of:

Clustering the said input images by single-link hierarchy clustering, where the image features are grouped into clusters with a properly chosen threshold, and a voting scheme is applied to select the best match during the evaluation.

10. Methods and apparatus for face recognition in the wild, comprising of, where Examining the relationship among the input images by clustering the said input images in the claim 8, further comprising of:

Clustering the said input images by using KNN (K-Nearest Neighbors) affinity matrix, where the image features can be updated by pseudo-cluster center with a properly chosen threshold, and the features belonging to the same class or identity will move closer after the updating.

11. Methods and apparatus for face recognition in the wild, comprising of, where Examining the relationship among the input images by clustering the said input images in the claim 8, further comprising of:

The said clustering in the invention can be supervised and learned clustering model.

12. Methods and apparatus for face recognition in the wild, comprising of, where Examining the relationship among the input images by clustering the said input images in the claim 8, further comprising of:

Managing a buffer to keep a number of the input images for performing the said clustering in real time application, where the said buffer size can be adjusted and dynamically changed according to different applications and use cases.

13.Methods and apparatus for face recognition in the wild, comprising of:

Extracting features from input images and testing images in a test dataset, where features can be learned via deep learning model, can be hand-crafted features;

Examining the relationship among the testing images in a testing database by clustering the said testing images;

Examining the relationship among the input images by clustering the said input images;

Adjusting and updating the image features according to the said cluster centers for both of the testing images and the input images;

Evaluating the similarity between the said input image and each of the testing image in the said database by calculating the distance of their feature sets;

Determining if they belong to the same class or same ID for the case of face recognition.