

Clustering Millions of Faces By Identity

Nisim Hurst

Friday 3 May 2019

Clustering Millions of Faces by Identity

The article was written by (Otto, Wang, and Jain 2018). It was cited 44 times according to Google Scholar. The task performed was face clustering. They used the F-measure metric over clusters with distractor images. They also developed their own metric for measuring internal cluster quality using just the k-top nearest neighbors.

Hypothesis

Deep features clustered using only the top-k nearest neighbors in rank-order clustering will produce a more scalable and a more accurate face clustering algorithm. This algorithm will be able to overcome the presence of millions distractor images and class imbalance.

The network architecture to produce a 320D feature vector was VGG16 proposed by (Simonyan and Zisserman 2014). The rank-order clustering algorithm is based on (Zhu, Wen, and Sun 2011). Their k-d tree implementation for calculating just the 200-top nearest neighbors is based on (Muja and Lowe 2014).

Evidence and Results

Evidence is presented first over a small dataset and then over an augmented version of the datasets with million of distractor images.

Dataset

For datasets were tested:

Table 1: Main characteristics of the four datasets that were used to test the improved CW.

| | # Instances | Resolution | Scenery | Author |
|---------------|---|----------------------------------|---|--|
| LFW | 13233 images of 5749. Only 1680 subjects have two or more photos. | ??, variable head angle | Color, different Poses and Backgrounds. | (Huang et al. 2008) |
| YTF | 3425 videos of 1595 subjects. | 100x100, variable enclosing area | Color, different Poses and Backgrounds. | (Wolf, Hassner, and Maoz 2011) |
| Webfaces | 123,654,141 distractor images. | 168x192 | Black and White, Frontal. | (Sanchez-Azofeifa, Caelli, and Cheng 2006) |
| CASIA-webface | 3200 of 126 subjects. | 120x165 | Color, Frontal. | (Mart'inez and Benavente 1998) |

Results

Contribution

Firstly, the authors improved the Rank-Order clustering algorithm proposed by (Zhu, Wen, and Sun 2011). The original Rank-Order has the disadvantage that it requires $O(n^2)$. The authors propose to use the FLANN library implementation of the randomized k-d tree algorithm to compute the list of top-k nearest neighbors. Just one iteration is used.

Secondly, the authors improved the internal quality metric of Modularization quality (MQ) (Mancoridis et al. 1998) by just counting shared neighbors in the top-k nearest neighbors list. Cluster's external quality was obviated.

Weaknesses

Future Work

References

- Huang, Gary B. et al. (Oct. 2008). "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments". In: *Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition*. Erik Learned-Miller and Andras Ferencz and Frédéric Jurie. Marseille, France. URL: <https://hal.inria.fr/inria-00321923>.
- Mancoridis, S. et al. (June 1998). "Using automatic clustering to produce high-level system organizations of source code". In: *Proceedings. 6th International Workshop on Program Comprehension. IWPC'98 (Cat. No. 98TB100242)*, pp. 45–52. DOI: [10.1109/WPC.1998.693283](https://doi.org/10.1109/WPC.1998.693283).
- Martínez, Aleix and Robert Benavente (June 1998). *The AR Face Database*. Tech. rep. 24. Cites in Scholar Google: <http://scholar.google.com/scholar?hl=en&lr=&client=firefox-a&cites=1504264687621469812>. Bel-latera: Computer Vision Center. URL: <http://www.cat.uab.cat/Public/Publications/1998/MaB1998>.
- Muja, M. and D. G. Lowe (Nov. 2014). "Scalable Nearest Neighbor Algorithms for High Dimensional Data". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36.11, pp. 2227–2240. ISSN: 0162-8828. DOI: [10.1109/TPAMI.2014.2321376](https://doi.org/10.1109/TPAMI.2014.2321376).
- Otto, C., D. Wang, and A. K. Jain (Feb. 2018). "Clustering Millions of Faces by Identity". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40.2, pp. 289–303. ISSN: 0162-8828. DOI: [10.1109/TPAMI.2017.2679100](https://doi.org/10.1109/TPAMI.2017.2679100).
- Sanchez-Azofeifa, A., T. Caelli, and L. Cheng (May 2006). "Component Optimization for Image Understanding: A Bayesian Approach". In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* 27.05, pp. 684–693. ISSN: 0162-8828. DOI: [10.1109/TPAMI.2006.92](https://doi.org/10.1109/TPAMI.2006.92).
- Simonyan, Karen and Andrew Zisserman (2014). "Very Deep Convolutional Networks for Large-Scale Image Recognition". In: pp. 1–14. ISSN: 09505849. DOI: [10.1016/j.infsof.2008.09.005](https://doi.org/10.1016/j.infsof.2008.09.005). arXiv: [1409.1556](https://arxiv.org/abs/1409.1556). URL: <http://arxiv.org/abs/1409.1556>.
- Wolf, L., T. Hassner, and I. Maoz (June 2011). "Face recognition in unconstrained videos with matched background similarity". In: *CVPR 2011*, pp. 529–534. DOI: [10.1109/CVPR.2011.5995566](https://doi.org/10.1109/CVPR.2011.5995566).
- Zhu, C., F. Wen, and J. Sun (June 2011). "A rank-order distance based clustering algorithm for face tagging". In: *CVPR 2011*, pp. 481–488. DOI: [10.1109/CVPR.2011.5995680](https://doi.org/10.1109/CVPR.2011.5995680).