

얼굴 인식을 통한 커스텀 데이터의 효율적 수집

*최희조 **김상준 *이유진 ***박경무 ***김정화 ***박구만
 *서울과학기술대학교 일반대학원 미디어IT공학과
 **서울과학기술대학교 나노IT디자인융합대학원 정보통신미디어공학전공
 ***서울과학기술대학교 전자미디어IT공학과
 *heejo0624@gmail.com

Collecting valid facial image dataset using face recognition

*Choi, Hee-jo **Kim, Sang-Joon
 *Lee, Yu-jin ***Park, Kyung-moo ***Kim, Chung-hwa ***Park,
 Goo-man
 *Dept. of Media It Engineering
 **Dept. of Information Technology and Media Engineering
 ***Dept. of Electronics and IT Media Engineering
 Seoul National University of Science and Technology

요약

인물에 대한 얼굴 이미지 데이터를 수집할 때 유효하지 않은 데이터를 수작업으로 걸러내는 것은 많은 시간과 인력의 투자를 필요로 한다. 얼굴 데이터 셋을 생성할 때 FaceNet을 거쳐 불필요한 정보들을 미리 걸러내고, 사람이 직접 얼굴 데이터의 유효성을 체크하는 수고를 덜어 얼굴 데이터 셋 생성에 있어서의 번거로움을 줄이고자한다. 본 논문에서는 FaceNet을 통해서 얼굴 이미지 데이터에 대한 데이터 수집 시, 더욱 좋은 성능으로 정제된 데이터 셋을 생성하고자 하였다.

1. 서론

딥러닝(Deep Learning)에 있어서 데이터의 양과 네트워크의 성능은 높은 상관관계가 있다. 따라서, 딥러닝 학습 결과의 정확도를 높이기 위해 많은 양의 데이터를 학습시킨다. 이 때, 대량의 데이터 셋을 모으는 방법으로 웹 크롤링이 자주 사용된다. 하지만, 기존의 방법으로는 불필요한 데이터도 높은 비율로 함께 수집되는 현상이 자주 발생한다. 이로 인해 수집된 데이터의 유효성에 대한 확인 및 분류 작업이 필요로 하며, 이를 위해 많은 시간과 인력이 필요하다는 문제점이 있다.

본 논문에서는 정제된 얼굴 데이터 수집 시 이미지를 저장하기 전에 FaceNet의 얼굴 분류기를 한번 거쳐 유효한 얼굴 영상 데이터 셋을 수집하도록 하는 시스템을 구축한다. 이 시스템에서는 데이터 수집 및 분류 시, 미리 훈련된 모델을 이용하여 추출기에 피드백을 제공하고, 이를 반영하여 데이터 셋을 생성하는데 소요되는 시간과 인력을 단축하고자 한다.

2. 본론

제안하는 얼굴 이미지 수집 시스템의 전체적인 과정은 그림 1과 같다. 먼저, 온라인 비디오 공유 플랫폼인 유튜브

(YouTube)에서 보이그룹 비투비 7명 멤버들에 대하여 총 200편의 동영상을 추출하였다. 수집한 모든 동영상에서 80 프레임마다 한 장씩 추출하여 총 120,000장의 이미지 데이터를 수집하였다. 수집된 이미지 중 소량의 이미지를 이용하여 클래스 별로 FaceNet의 분류기 모델에 미리 학습시킨다. 그 후, 수집된 전체 이미지 데이터 셋의 얼굴 검출 및 정렬을 위해 FaceNet의 MTCNN[2] (Multi Task Cascaded Convolutional Neural Network)을 사용한다. 이 과정에서 그림 2의 왼쪽 예시와 같이 얼굴이 검출되지 않은 데이터는 유효하지 않은 데이터로 판단하여 반환한다. 그림 2의 중간 그림과 같이 2명 이상의 얼굴이 검출되는 데이터 또한 효용가치가 떨어지는 데이터로 판단하여 반환한다. 그리고 그림 2의 오른쪽 그림처럼 얼굴이 올바르게 정렬된 데이터들은 그림 3과 같이 클래스 별로 정렬 및 분류된다. 이 후, 전체 데이터에 대해 분류기를 학습시켜 얼굴 인식 및 얼굴 클러스터링을 진행한다. 검출기에 미리 학습된 분류기의 정보를 탑재하여 얼굴 클러스터링에 있어서 가이드라인으로서의 역할을 한다. 이는 얼굴 클러스터링에 대한 label이 없기 때문에 클러스터링에 대한 정확도를 판단할 수 없는 부분을 보완하여 작업자의 편의를 제공할 것으로 예상된다.

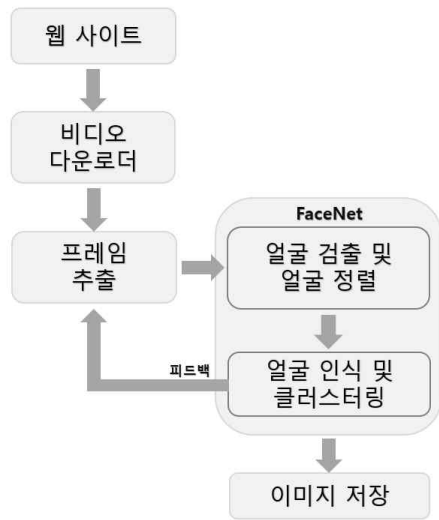


그림 1. FaceNet을 이용한 얼굴 데이터 수집 시스템



그림 2. 이미지 데이터의 얼굴을 정렬(Face alignment)한 것을 표현해놓은 모형. 유효하지 않은 데이터(왼쪽, 중간)[4],[5]와 얼굴 검출이 올바르게 된 데이터(오른쪽)[6]의 예시.

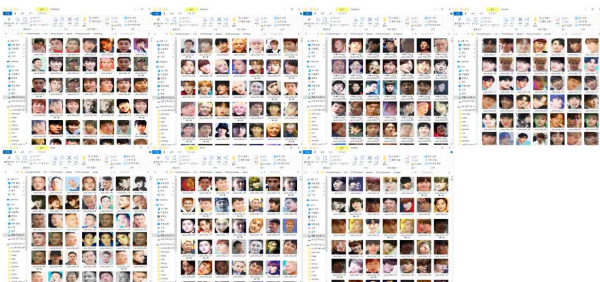


그림 3. 7개 클래스에 대해 얼굴 검출 및 정렬을 하여 데이터를 수집한 폴더의 모습

3. 결론

본 논문에서 122,271장의 데이터를 FaceNet 아키텍처에 통과시킨 결과를 표 1에 나타내었다. FaceNet을 사용하기 전과 후의 결과를 비교하여 29,929장의 불필요한 데이터 셋을 필터링함으로써 이전보다 수월하게 데이터 셋을 확보할 수 있다는 것을 증명하였다. 또한, 표 2와 같이 기존에 10,000장의 정제되지 않은 데이터 셋을 정리하는데 클래스 당 평균 약 15.9시간이 들었던 반면에, 위의 시스템을 통해 평균 약 14.3시간으로 시간이 단축되는 것

표 1. FaceNet의 얼굴검출 및 정렬을 거치기 이전과 이후의 데이터 셋(단위 : 장)

FaceNet 전의 데이터 셋	FaceNet 후의 데이터 셋	유효하지 않은 데이터 셋
122,271	92,342	29,929

표 2. 10,000장당 얼굴 데이터 셋 수집 및 분류작업에 걸리는 평균시간 측정(단위 : hour)

	웹 크롤링 데이터 수집 평균 시간(per class)	분류 작업 평균 시간(per class)
기존 방식		15.9h
FaceNet을 거친 방식	6.2h	14.3h

이 확인된다. 따라서, 학습을 위한 데이터를 만드는 작업에서의 오류율도 크게 줄어들 것으로 예상된다.

향후 계획으로는 본 논문의 시스템을 자동화하여 데이터 셋 생성의 효율성을 높이고, 시스템의 성능을 개선할 예정이다. 이를 위해 데이터 증강(Data augmentation) 등의 방법을 통해 데이터 셋의 양을 더욱 늘려서 시스템의 유효성에 대하여 실험할 예정이며, 확보해놓은 데이터들을 이용해 FaceNet의 실시간 비디오 분량체크기를 만들어 동영상 내의 등장인물의 얼굴을 실시간으로 인식하여 동영상 내 총 등장시간을 확인할 수 있는 프로그램을 개발 중이고, 이에 따른 얼굴 검출 및 얼굴 인식의 정확도를 높이는 연구를 진행할 예정이다.

Acknowledgement

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2017-0-00217, 투명도와 레이어 가변형 실감 사이니지 기술 연구)

참고문헌

- [1] Florian Schroff, Dmitry Kalenichenko, James Philbin, FaceNet : A Unified Embedding for Face Recognition and Clustering, IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2015, June, 2015.
- [2] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, Yu Qiao, Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks, IEEE, April, 2016.
- [3] D. Chen, S. Ren, Y. Wei, X. Cao, and J. Sun. Joint cascade face detection and alignment. In Proc. ECCV, July, 2014.
- [4] 더셀럽, [더셀럽포토]비투비 정일훈, '잘 다녀와' [Internet] http://theceluv.com/upimages/gisaimg/202005/1589175655_321605.jpg
- [5] 서은광, '자', [Internet] <http://www.instagram.com/p/BmbRw2XnMO6/>
- [6] 옥성재, '뭐지? 뭐지?', [Internet] <http://www.instagram.com/p/B5nRlFpBK3p/>