

K-Means Clustering 기반 온디바이스를 위한 효율적인 얼굴 인식 알고리즘

조현준¹, 양병찬², 차주형³, 권용인⁴¹배재대학교 전자공학과 학부생²경북대학교 전자공학부 학부생³과학기술연합대학원대학교 인공지능학과⁴한국전자통신연구원 온디바이스시스템SW연구실

(lckc9811, cks27112)@gmail.com,

(jh.cha, yongin.kwon)@etri.re.kr

Resource-Efficient On-Device Face Recognition Using K-Means Clustering

Hyunjun Cho¹, Byeongchan Yang², Joohyung Cha³, Yongin Kwon⁴¹Dept. of Electronic Engineering, Paichai University²Dept. of Electrical Engineering, Kyungpook National University³Artificial Intelligence, University of Science and Technology⁴On-Device System SW Laboratory, Electronics and Telecommunications Research Institute

요 약

본 논문은 안면 인식 시스템에서 아크페이스 기반 임베딩 추출 기법의 성능을 K-means clustering 알고리즘과 결합하여 검증한 연구이다. 기존의 아크페이스 기반 안면 인식 모델은 개별 이미지 단위의 특징 벡터를 생성하므로, 데이터셋 내 다각도(정면, 좌측, 우측, 하향, 상향) 안면 이미지를 직접적으로 반영하지 못하는 한계가 존재한다. 이를 개선하기 위해, 본 연구에서는 동일 인물의 다수 임베딩을 K-means clustering으로 군집화하여 대표 임베딩을 생성한다. 이 기법으로 임의의 이미지와 유사도 비교를 진행했을 때 인식 성능 개선을 확인했다.

1. 서론

최근 클라우드 기반 얼굴 인식 시스템에서는 사용자의 얼굴 이미지를 다수 저장한 후, 각 이미지에서 추출한 임베딩을 활용하여 비교 연산을 수행한다. 예를 들어, 한 명의 사용자를 등록하는 과정에서 수백 장의 얼굴 이미지를 촬영하여 저장하는 경우가 일반적이다. 그러나 실시간 환경에서 새로운 얼굴 임베딩을 비교할 때에는 저장된 임베딩의 총 개수(사용자 수 × 이미지 수)만큼 연산이 필요하다. 특히 연속 촬영된 이미지들은 유사하거나 중복된 정보를 포함하는 경우가 많아 서버의 연산 자원과 저장 공간 낭비를 초래한다.

이를 해결하기 위해 본 연구에서는 K-means clustering [2] 기법을 적용하여 각 사용자별 임베딩 집합으로부터 대표 임베딩을 추출하는 방식을 사용한다. K-means clustering은 계산 구조가 단순하고 군집 중심을 통해 직관적인 대표성을 제공하므로, 온디바이스 환경에서도 효율적으로 동작할 수 있다는 장점이 있다. 구체적으로, 새로운 임베딩이 추가될 때마다 군집의 중심점을 갱신하는 과정을 반복하여 안정적인 대표 벡터를 형성한다. 이후 비교 단계에서는 전체 임베딩이 아닌 대표 임베딩만을 활용함으로써 연산량을 대폭 줄일 수 있다. 예를 들어, 원본 데이터가

사용자당 500장의 이미지로 구성된 경우, K-means clustering 기법을 통해 정면, 상, 하, 좌 등 네 가지 대표 임베딩을 추출한다면 비교 연산량은 약 1/100 수준으로 감소한다.

본 연구는 이러한 접근 방식을 통해 실시간 얼굴 인식 환경에서 연산 시간을 효과적으로 단축하고, 단일 이미지 및 영상 기반 실험을 통해 인식 성능 향상 가능성을 검증하고자 한다.

2. 실험 환경

<표 1> 시스템 구성 요약

구분	항목	구성
H/W	통합 제어 모듈	라즈베리파이 5
	가속기	Hailo-8L
	카메라 모듈	NOIR 8MP Sony IMX219
S/W	운영체제	Raspberry Pi OS, Linux
	사용 모델	MobileFaceNet, Arcface loss

표 1은 본 연구에 사용된 하드웨어 및 소프트웨어 환경을 나타낸다. 제어 장치로는 라즈베리파이5를 사용하였으며 M.2 HAT+ 보드를 활용하여 Hailo-8L AI 가속기

모듈과 연결하였다. 임베딩 추출에는 MobileFaceNet을 백본으로 하고 아크페이스[1]를 손실 함수로 학습시킨 MobileFaceNet 모델을 사용했다. 해당 모델로 사전 추출한 임베딩을 K-means clustering 알고리즘의 적용 여부에 따라 비교하여 각각의 인식률을 측정하였다.

3. 실험 결과

3.1 새로운 이미지와의 임베딩 비교

그림 1에서 (1)과 (2)는 각각 단일 임베딩 추출 시 사용된 단일 이미지와 K-means clustering 기법으로 추출 시 사용된 이미지이다.



(1) 단일 이미지 예시



(2) 추출된 대표 이미지 예시 (왼쪽부터 좌, 위, 정면, 아래)
(그림 1) K-평균 군집화를 통해 추출한 이미지의 결과

표 2는 위 (1)과 (2)에서 추출한 임베딩을 새로운 이미지(정면, 우, 좌, 위, 아래)와 비교했을 때의 유사도를 나타낸 것이다. 본 실험에서는 안면 각도별 5개 방향(Front/Right/Left/Up/Down)에 대해 코사인 유사도를 지표로 비교하여 개선율을 나타냈다.

실험 결과, 임베딩 수가 많아질수록 유사도가 높아지는 경향을 확인했다. 또한, 단일 임베딩에서 K=4에서의 개선율은 평균 18.1%로 크게 향상되었다. 반면, K=4에서보다 K=500으로 임베딩 개수를 125배 늘려도 유사도 차이는 평균 1.76%에 그쳤다. 이는 소수 대표 임베딩도 다각도 인식에서 높은 유사도를 유지할 수 있음을 시사한다.

<표 2> k값에 따른 유사도 비교

	Front	Right	Left	Up	Down
Single Embedding (front)	0.8300	0.7851	0.7586	0.5712	0.7431
K=4 (front,right,up,down)	0.8323 (+2.7%)	0.9381 (+19.4%)	0.8377 (+10.4%)	0.8218 (+43.8%)	0.8489 (+14.2%)
500 embedding (front,right,left,up,down)	0.8672 (+7%)	0.9390 (+19.6%)	0.8447 (+11.3%)	0.8219 (+43.8%)	0.8807 (+18.5%)

3.2 연산시간 단축 비교

표 3은 3.1에서 했던 실험에서 많은 임베딩을 사용하는 것이 결코 좋은 선택지가 아닌 이유를 보여주는 실험 결과이다.

임베딩 개수를 많이 했을 때, 유사도가 가장 높았지만 그만큼 연산 리소스를 잡아먹는다. 이는 온디바이스 환경에서 인식할 수 있는 사람을 늘릴수록 인식 지연이 발생

할 수 있다. K-means clustering 기법은 500개의 임베딩을 사용하여 비교했을 때와 유사도는 비슷하지만 평균적으로 27.5배 더 짧은 인식 시간을 보여준다.

<표 3> k값에 따른 연산시간 비교

	Saved Embedding Load Time (ms)	Similarity Comparison Time (ms)
Single embedding(front)	0.5	0.4
K=4	0.5 (x1)	0.5 (x1.25)
500 embedding	9.9 (x19.8)	14.1 (x28.2)

4. 결론 및 향후 계획

본 연구는 온디바이스 환경의 연산 및 메모리 제약 조건 하에서 아크페이스 기반 임베딩(MobileFaceNet 백본)과 K-means clustering 기반 대표 임베딩 추출을 결합함으로써, 비교 대상 임베딩 수를 효과적으로 축소하면서도 다각도 얼굴 인식의 견고성을 확보할 수 있음을 실험적으로 검증하였다. 단일 임베딩과 500개의 임베딩, K-means clustering 기법(K=4)을 적용한 대표 임베딩을 사용하는 총 3가지 방법으로 임의의 이미지와 비교했을 때의 유사도와 연산 시간을 측정하였다. 각 방법은 장단점을 가지고 있지만 K-means clustering 방식은 나머지 두 방식의 장점을 융합한 결과를 보여주었다. 단일 임베딩만을 활용했을 때 발생하는 비정면 각도에서의 유사도 저하와 500개의 임베딩을 활용했을 때 발생하는 연산 시간 지연을 대표 임베딩(정면, 상향, 하향, 좌측)으로 보완함으로써, 연산량 절감과 유사도 개선을 동시에 달성할 수 있음을 확인하였다.

향후 연구에서는 다양한 조건(조명 변화, 얼굴 가림 등)을 포함한 대규모 데이터셋을 활용하여 본 방법론의 일반화 능력을 체계적으로 검증할 계획이다. 더 나아가, 실사용 환경에서의 안정성을 확보하기 위해 클러스터별 적응형 임계값 설정, 임베딩 품질 기반 가중치 도입과 같은 정교한 전략을 적용할 예정이다. 또한, 얼굴 검출과 임베딩 추출을 이중 Hailo 가속기에서 병렬화하여 실시간 인식 성능을 한층 강화하는 방안을 모색하고자 한다.

사사

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구 결과임 (No.RS-2024-00459797, No.RS-2023-00277060, No.RS-2025-02217404, No.RS-2025-02214497, No.RS-2025-02216517)

참고문헌

- [1] Jiankang Deng, Jia Guo, Jing Yang, Niannan Xue, Irene Kotsia, Stefanos Zafeiriou, ArcFace: Additive Angular Margin Loss for Deep Face Recognition
- [2] Yiqun T. Chen, Daniela M. Witten, Selective inference for k-means clustering