

논문 2020-2-7 <http://dx.doi.org/10.29056/jsav.2020.12.07>

딥러닝 기반의 얼굴인증 시스템 설계 및 구현

이승익*†

Design and Implementation of a Face Authentication System

Seungik Lee*†

요 약

본 논문에서는 딥러닝 프레임워크 기반의 얼굴인증 시스템에 대하여 제안한다. 제안 시스템은 딥러닝 알고리즘을 활용하여 얼굴영역 검출과 얼굴 특징 추출을 수행하고, 결합베이지안 학습 모델을 이용하여 얼굴인증을 수행한다. 제안 얼굴인증 알고리즘에 대한 성능 평가는 다양한 얼굴 사진들로 구성된 데이터베이스를 이용하여 수행하였으며, 한 명에 대한 얼굴 영상은 2장으로 구성하였다. 또한 얼굴인증 실험은 딥 뉴럴 네트워크를 통한 2048차원의 특징과 그 유사성을 측정하기 위해 결합베이지안 알고리즘을 적용하였으며, 얼굴인증에 실패한 동일오율을 계산함으로써 성능평가를 수행하였다. 실험 결과, 딥러닝 특징과 결합베이지안 알고리즘을 사용한 제안 방법은 1.2%의 동일오율을 보였다.

Abstract

This paper proposes a face authentication system based on deep learning framework. The proposed system is consisted of face region detection and feature extraction using deep learning algorithm, and performed the face authentication using joint-bayesian matrix learning algorithm. The performance of proposed paper is evaluated by various face database, and the face image of one person consists of 2 images. The face authentication algorithm was performed by measuring similarity by applying 2048 dimension characteristic and combined Bayesian algorithm through Deep Neural network and calculating the same error rate that failed face certification. The result of proposed paper shows that the proposed system using deep learning and joint bayesian algorithms showed the equal error rate of 1.2%, and have a good performance compared to previous approach.

한글키워드 : 딥러닝, 얼굴인식, 영상처리, 결합베이지안, 능동적 형태 모델

keywords : deep learning, face authentication, image processing, joint bayesian, active shape model

1. 서론

최근 딥러닝 기술의 발전과 함께 다양한 분야

에 딥러닝을 기법을 적용하려는 연구가 등장하고 있다. 딥러닝 기술은 보다 더 잘 설계된 네트워크를 학습함으로써, 기존보다 더 좋은 특징 벡터를 추출하고 더 정확한 분류가 가능하다. 얼굴 인증 분야 또한 딥러닝 기술을 적용하여 제한된 환경에서 인간보다 더 좋은 성능을 보이고 있다 [1]. 이에, 본 논문에서는 딥러닝 프레임워크에

* 연암공과대학교 스마트소프트웨어학과

† 교신저자: 이승익(email: seungik@yc.ac.kr)

접수일자: 2020.08.10. 심사완료: 2020.08.24.

게재확정: 2020.12.21.

기반한 얼굴인증 시스템에 대하여 제안한다. 본 논문은 기존의 연구결과[2]를 바탕으로 보다 강한 얼굴인증 시스템을 구현하기 위하여, 딥러닝 기법을 얼굴 검출과 특징 추출 부분을 적용함으로써, 기존 연구결과보다 보다 더 정확도가 높은 시스템을 구현하였다. 보다 구체적으로 기존 연구[2]는 얼굴 영역 검출을 위해 AdaBoost[3] 알고리즘을 사용한 반면, 본 연구에서는 딥 네트워크 기반의 YOLO(You Only Look Once)[4] 알고리즘을 적용하였고, 특징추출 모듈을 기존연구에서는 고차원 국부이진패턴(HD-LBP)[5]를 사용한 반면, 본 연구에서는 딥 네트워크에 기반한 수정된 Resnet50 [6]을 적용하였다.

최종적으로 제안 시스템은 얼굴영역 검출, 특징추출, 그리고 인증의 단계로 구성되며, 정밀 얼굴 검출은 YOLO 알고리즘과 능동형태모델(Active Shape Model, ASM)[7]을 이용하여 특징점(Landmark)을 검출하고, 정규화 함으로써 얼굴영역을 획득하였다. 얼굴인증에 필요한 특징추출 단계는 성능향상에 많은 영향을 주며, 기존 연구의 HD-LBP 대신에 Resnet50 딥 네트워크를 사용하여 얼굴 특징을 추출하였다. 마지막으로 얼굴 인증은 추출된 특징을 이용하여 결합베이지안[8] 학습 모델을 생성하고, 테스트 단계에서 두 얼굴 영상의 유사도를 비교함으로써 수행된다. 제안하는 얼굴인증 방법의 성능평가는 500명으로 이루어진 얼굴 데이터베이스를 이용하여 수행하였다.

본 논문의 구성은 서론에 이어 II절에서는 얼굴영역 획득, 특징 추출 및 인증 등 제안 방법에 대하여 자세히 소개한다. III절에서는 제안 시스템의 실험 결과를 소개하며, 마지막으로 IV절에 본 논문의 결론을 제시한다.

2. 제안방법

본 논문에서 제안하는 딥러닝 프레임워크 기반의 얼굴인증 시스템 블록도는 그림 1과 같다. 먼저, 입력영상에서 YOLO객체 검출 방법과 ASM 알고리즘으로 얼굴영역을 검출하고 정합한다. 다음으로 Resnet50 딥 네트워크를 이용하여 얼굴의 특징을 추출한 후, 최종적으로 사전 학습된 결합베이지안 모델과 유사도를 비교함으로써 입력 영상이 본인 또는 타인인지의 여부를 판단한다.

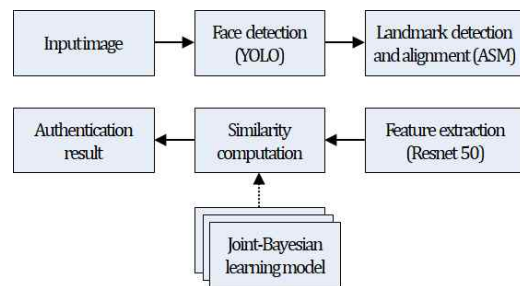


그림 1. 제안한 방법의 시스템블록도
Fig. 1. A system block diagram for the proposed method.

2.1 얼굴영역 획득

입력 영상에서 얼굴 영역 검출과정은 얼굴인증을 위한 첫 번째 단계이다. 본 논문에서는 얼굴 검출을 위하여 딥 네트워크에 기반한 객체 검출 방식인 Fast R-CNN[9], Faster R-CNN[10], SSD(Single Shot MultiBox Detector)[11], YOLO[4] 등 다양한 방법들 중에서, 현재 객체 탐지 속도가 가장 우수한 YOLO 알고리즘을 사용하였다. YOLO는 실시간으로 다수의 객체를 검출하는 알고리즘으로서, 입력 영상에서 객체의 영역과 종류를 동시에 검출/분류한다. YOLO는 입력 영상을 격자 셀로 분할하여, 각 셀 별로 객체 존재 유무와 종류에 대한 확률 값을 계산하고, 확률값과 NMS (Non-Maximum

Suppression)을 사용하여 객체를 검출한다. 본 논문에서는 얼굴 객체만을 검출하므로, 다양한 YOLO 모델 중에서 적은 수의 객체 검출에 적합한 비교적 가벼운 모델인 Tiny 모델을 사용하여 얼굴을 검출하였다. 본 논문에서 사용한 Tiny YOLO 모델의 구조는 그림 2와 같다.

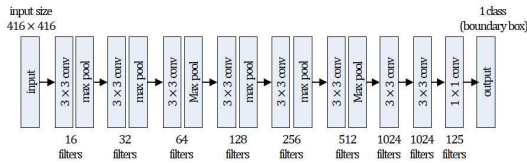


그림 2. tiny YOLO모델의 구조
Fig. 2. A structure of tiny YOLO model.

Tiny YOLO 모델을 이용하여 얼굴 검출 후, 보다 정밀한 얼굴 영역 획득을 위하여, 본 논문에서는 능동형태모델을 이용하여 랜드마크를 추출하고, 랜드마크 정보를 이용하여 얼굴을 정규화하였다. 능동형태모델의 특징은 통계적인 모델에 기반한 특징점을 추출하는 방법으로 검출 속도가 빠르며, 모델의 변형된 형태를 표현이 용이하다. 그림 3은 능동형태모델을 이용한 얼굴 특징점 및 그 위치를 9개의 점으로 나타내었다.

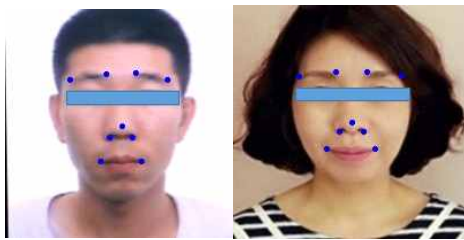


그림 3. ASM-based face landmark 추출
Fig. 3. ASM-based face landmark extraction.

2.2 인증을 위한 특징 추출

본 논문에서는 얼굴의 특징 파라미터로 CNN

구조의 딥러닝 알고리즘을 사용하였다. 일반적으로 CNN은 합성곱과 풀링(Pooling)을 순차적으로 반복하며 다단계로 구성한 후, 마지막단에 다층 퍼셉트론 신경망을 구성하여 활용한다. 합성곱은 영상에서의 국부적특징을 추출하고, 풀링은 입력 차원을 감소시키는 과정이다. 본 논문에서는 얼굴인증을 위하여 분류를 위한 마지막 레이어인 Fully-connected 이전의 레이어 출력을 얼굴인증의 특징 파라미터로 이용하였다. 제한한 논문에서 사용한 CNN 구조는, 그림 4와 같이 50개의 레이어로 이루어진 Residual Network를 수정하여 사용하였고, 마지막 레이어의 전 레이어인 2048차원의 특징을 추출하고, 이를 얼굴 파라미터로 사용하였다. CNN 네트워크의 학습은 이미 지넷으로 사전 학습된 모델을 가져와, 테스트 데이터셋 외의 500명에 대하여 추가로 구성한 데이터를 이용하여, 마지막 Fully-connected 레이어를 500명을 분류하는 네트워크로 구조를 변경하고, 모델을 재학습함으로써 모델 파라미터를 최적화 하였다. 최종적으로 얼굴 인증을 위하여 사용하는 특징은 2048차원으로 구성되며, 이는 다음 단계인 결합 베이지안 모델의 입력으로 사용된다.

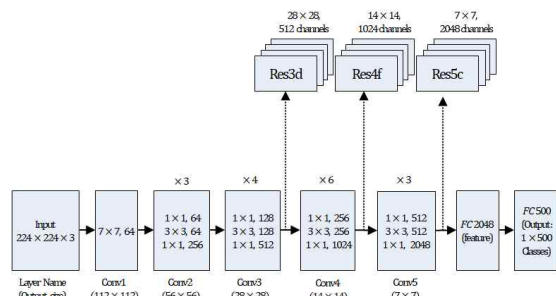


그림 4. resnet 50 네트워크 구조
Fig. 4. Structure of resnet 50 network.

마지막으로, 두 얼굴의 유사도 비교는 CNN 특징과 결합베이지안 알고리즘으로 수행된다. 결합

베이지안 모델은 입력얼굴의 내부의 분석학적인 특징과 눈으로 보이는 특성 모두를 잘 표현해주는 모델이다. 그림 5에 보는 바와 같이 분석학적인 특성은 공통적으로 얼굴이 가지는 특성을 나타내고, 눈으로 보이는 특성은 사람들을 구분할 수 있는 특성을 나타내고 있다.

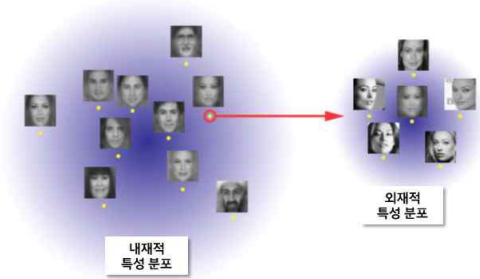


그림 5. 얼굴 표현
Fig. 5. face representation

얼굴 특징 벡터 x 는 분석학적인 특성 μ 와 외향적 특성 ϵ 의 합으로 표현될 수 있다.

$$x \approx \mu + \epsilon, \tag{1}$$

위 식(1)에서 μ 와 ϵ 은 독립 정규분포특성을 가지며, 각각의 특성은 $\mu \sim N(0, S_\mu)$, $\epsilon \sim N(0, S_\epsilon)$ 와 같은 공분산을 가진다고 가정하였다. 또한, x 는 식 (2)와 같이 표현된다.

$$x \sim S_\mu + S_\epsilon. \tag{2}$$

두 얼굴의 유사도 계산을 위한 특징벡터 x_1 과 x_2 에 대한 유사도 r 은 식(3)과 같은 조건부 결합확률밀도함수의 비를 이용하여 구할 수 있다.

$$r(x_1, x_2) = \log \frac{P(x_1, x_2 | H_I)}{P(x_1, x_2 | H_E)}, \tag{3}$$

여기서 H_I 는 두 얼굴 x_1 과 x_2 는 같은 얼굴이라는 가정을 나타내며, H_E 는 x_1 과 x_2 은 서로 다른 얼굴의 가정임을 나타낸다.

3. 실험 및 결과

본 논문에서 제안하는 얼굴인증 시스템의 성능평가는 기존연구[2]와의 비교를 위하여 동일한 조건으로 수행하였다. 성능평가는 한 명에 대하여 2장(학습 및 테스트용)으로 구성된 500명의 얼굴 영상 데이터베이스를 이용하였다. 그림 6에서는 본 논문에서 사용된 얼굴 샘플영상을 나타내며, 기존연구에서와는 다른 샘플 영상을 나타내었다. 또한, 얼굴 부분에서의 눈 영역은 모두 마스트 처리하여 영상을 나타내었다.

얼굴인증 실험은 딥 뉴럴 네트워크를 통한 2048차원의 특징과 유사도를 측정하기 위한 결합 베이지안 알고리즘을 적용하였으며, 동일오율(equal error rate, EER)을 통한 얼굴인증 실패율을 이용하여 본 논문의 성능평가를 수행하였다. 최종적으로 본 논문에서 제안하는 딥러닝 특징과 결합베이지안 알고리즘을 사용한 얼굴 인증 시스템은 1.2%의 동일오율을 보였다. 기존연구에서 고차원 국부이진패턴 특징과 유클리디언 거리를 사용할 경우, 동일오율이 19.2%, 결합베이지안 알고리즘을 사용할 경우에는 2.3%를 보였음에 반하여, 제안 시스템의 성능은 1.2%임을 확인하였으며, 이로부터 제안 시스템이 1.1%의 성능향상을 실험을 통해 확인하였다.

더불어, 본 논문에서 제안하는 시스템의 효용성을 확인하기 위하여 실 환경에서 시스템을 구현하였으며, 실행 예시를 그림 6에 나타내었다. 그림 7(a)는 학습에 이용한 신분증 영상을 보이고 있으며, 그림 7(b), 7(c), 7(d)는 가려짐, 회전 등의 변화를 가한 영상을 입력으로 인증여부를 수행하는 화면을 보여준다. 학습영상의 품질이 비교적 선명하지 못하지만, 테스트 영상을 각각 입력하였을 경우, 각각의 유사도 값은 64%, 54%, 63%를 보였다. 더불어 제안 시스템은 동일오율을 계산함에 있어 임계값을 51%로 설정하였고,

임계값 이상이면 동일인, 이하이면 타인으로 판단하는 로직을 설정하였다. 이러한 제안 시스템 설정하에 상기 실 환경 실험은 그림 7(b), 7(c), 7(d)와 같은 입력의 모든 경우에 대하여 본인으로 판단하는 것을 확인할 수 있었다.

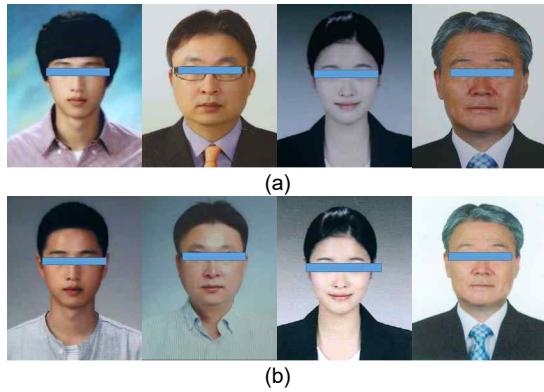


그림 6. 얼굴 이미지 예제
 (a) 훈련 이미지 (b) 시험 이미지
 Fig. 6. The examples of face images
 (a) training images (b) test images.

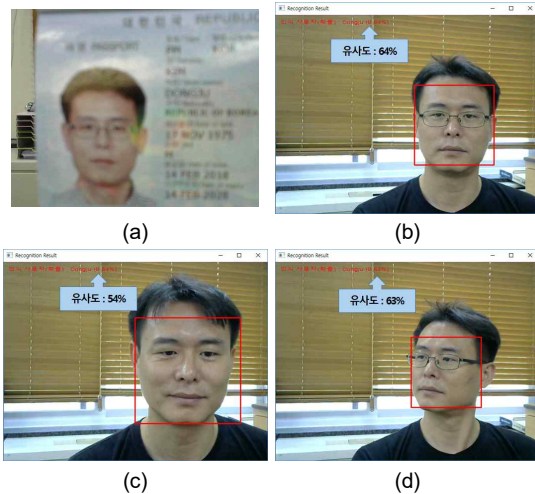


그림 7. 제안한 시스템의 결과 예시
 (a) 원본 (b) 정면얼굴 (c) 안경없는 얼굴
 (d) 여러각도의 이미지
 Fig. 7. An example of proposed system
 (a) original image (b) frontal image
 (c) non-glasses image (d) pose-variation image.

4. 결론

본 논문에서는 딥 네트워크 프레임워크 기반의 얼굴인증 시스템을 제안하였다. 제안 시스템은 딥러닝 알고리즘을 얼굴 검출 및 특징 추출에 적용함으로써, 기존 연구 결과를 개선하였으며, 이를 실험으로부터 증명하였다. 실제, 500명으로 구성된 얼굴 데이터베이스를 이용한 동일 조건의 실험으로부터 기존의 HD-LBP 특징 추출에 기반한 방법이 2.3%의 동일오율을 보임 반면, 제안 시스템은 1.2%의 동일오율을 보였다. 이로부터, 제안하는 얼굴인증 시스템이 딥 네트워크 기법을 적용함으로써, 우수한 얼굴인증 성능을 보여줌을 확인하였다. 향후, GPU 연산이 지원되는 임베디드 시스템 상에 실시간으로 동작 할 수 있도록, 본 연구 결과의 최적화와 구현에 관한 연구를 수행할 계획이다.

참고 문헌

- [1] Y. Taigman, M. Yang, M. A. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification", in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1701-1708, 2014. <https://doi.org/10.1109/CVPR.2014.220>
- [2] D.Kim, S.Lee, S.Kang, "Implementation of a Face Authentication Embedded System Using High-dimensional Local Binary Pattern Descriptor and Joint Bayesian Algorithm", Journal of the Korea Institute of Information and Communication Engineering, vol. 21, no. 9, pp. 1674-1680, 2017. <https://doi.org/10.6109/jkiice.2017.21.9.1674>
- [3] W. Yu, J. Xiu, C. Liu, and Z. Yang, "A depth cascade face detection algorithm

- based on adaboost”, in Proceeding of IEEE International Conference on Network Infrastructure and Digital Content, Beijing, China, pp. 103-107, Sep. 2016. <https://doi.org/10.1109/ICNIDC.2016.7974544>
- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection”, Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788, 2016. <https://doi.org/10.1109/CVPR.2016.91>
- [5] D. Chen, X. Cao, F. Wen, and J. Sun, “Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification”, in Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, pp. 3025-3032, June 2013. <https://doi.org/10.1109/CVPR.2013.389>
- [6] S. Zagoruyko, and N. Komodakis. “Wide residual networks”, arXiv preprint arXiv:1605.07146, 2016. <https://arxiv.org/abs/1605.07146>
- [7] K.-M. Jenog and J.-H. Kim, “Face classification and analysis based on geometrical feature of face”, Journal of the Korea Institute of Information and Communication Engineering, vol. 16, no. 7, pp. 1495-1504, July 2012. <https://doi.org/10.6109/jkiice.2012.16.7.1495>
- [8] D. Chen, X. Cao, L. Wang, F. Wen, and J. Sun, “Bayesian face revisited: A joint formulation”, in Proceeding of European Conference on Computer Vision, Firenze, Italy, pp. 566-579, Oct. 2012. https://doi.org/10.1007/978-3-642-33712-3_41
- [9] R. Girshick, “Fast r-cnn”, ICCV, pp.1440-1448, 2015. https://openaccess.thecvf.com/content_iccv_2015/papers/Girshick_Fast_R-CNN_ICCV_2015_paper.pdf
- [10] S. Ren, K. He, R. Girshick, and J. Sun, “Faster rcnn: Towards real-time object detection with region proposal networks”, NIPS, pp.91-99, 2015. <https://doi.org/10.1109/TPAMI.2016.2577031>
- [11] W Liu, D Anguelov, D Erhan, C Szegedy, “SSD: Single Shot MultiBox Detector”, European Conference on Computer Vision, pp. 21-37, 2016. <https://arxiv.org/pdf/1512.02325.pdf%EF%BC%89>

저자 소개



이승익(Seungik Lee)

1994.2 경북대학교 전자공학과 학사 졸업
 2002.8 콜로라도주립대학교 전기컴퓨터공학 석사 졸업
 2005.8 경북대학교 전자공학과 박사 졸업
 2008.4-현재 : 연암공과대학교 스마트소프트웨어학과 교수
 <주관심분야> 머신러닝, 얼굴인식