

ResNet 모델을 이용한 눈 주변 영역의 특징 추출 및 개인 인증

김민기[†]

Feature Extraction on a Periocular Region and Person Authentication Using a ResNet Model

Min-Ki Kim[†]

ABSTRACT

Deep learning approach based on convolution neural network (CNN) has extensively studied in the field of computer vision. However, periocular feature extraction using CNN was not well studied because it is practically impossible to collect large volume of biometric data. This study uses the ResNet model which was trained with the ImageNet dataset. To overcome the problem of insufficient training data, we focused on the training of multi-layer perception (MLP) having simple structure rather than training the CNN having complex structure. It first extracts features using the pretrained ResNet model and reduces the feature dimension by principle component analysis (PCA), then trains a MLP classifier. Experimental results with the public periocular dataset UBIPr show that the proposed method is effective in person authentication using periocular region. Especially it has the advantage which can be directly applied for other biometric traits.

Key words: Periocular Region, Authentication, CNN, MLP

1. 서론

개인의 고유한 생체정보(biometric trait)를 이용하여 개인을 식별하는 생체인식 기술은 생체정보의 다양성 확대와 함께 그 응용 분야도 지속적으로 확대되고 있다. 오래전부터 범죄 수사에 이용되었던 지문은 물론 홍채, 정맥, 장문, 얼굴 등의 생체정보를 활용한 온라인 인증 시스템, 근대 관리 시스템, 공항 출입국 관리 시스템 등이 보급되어 사용되고 있다. 최근에는 CCTV를 활용한 범죄 수사가 확대되고, 줌인(zoom in) 기능을 갖춘 고해상도 CCTV가 보급되면서 얼굴 인식 기술을 이용하여 보행자를 자동으로 식별하기 위한 연구들이 이루어지고 있다[1].

생체인식 시스템은 크게 생체정보 획득 단계와 생체정보 식별 단계로 나누어 볼 수 있다. 생체정보 획득 단계에서는 인식 대상자의 생체정보를 수집하여 개인을 식별할 수 있는 특징을 추출한 후 생체정보 데이터베이스에 저장한다. 생체정보 인식 단계에서는 인식 시점에 획득한 생체정보로부터 특징을 추출한 후 이를 생체정보 데이터베이스에 저장된 특징과 비교하여 특정인을 식별하게 된다. 지문은 접촉식 방식에 의하여 생체정보의 획득이 이루어지며, 홍채와 장문은 피사계 심도가 낮은 가까운 거리에서 획득되기 때문에 생체정보를 획득하기 위해서는 인식 대상자의 협조가 필수적이다. 이와 달리 얼굴은 먼 거리에서도 획득이 가능하기 때문에 상대적으로 인식 대

* Corresponding Author : Min-Ki Kim, Address: (660-701) Jinju-daero 501, Jinju, Gyeongnam, Korea, TEL : +82-55-772-1393, FAX : +82-55-772-1389, E-mail : mkkim@gnu.ac.kr

Receipt date : Oct. 11, 2019, Revision date : Nov. 9, 2019
Approval date : Nov. 18, 2019

[†] Dept. of Computer Engineering, Gyeongsang National University Engineering Research Institute

상자의 협조 없이도 생체정보 획득이 가능하다. 본 연구에서는 개인 식별을 위하여 얼굴 중 눈 주변 영역(periorcular region)에 나타나는 생체정보를 이용한다. 눈 주변 영역에 대한 생체정보는 원거리에서 획득이 가능하기 때문에 인식 대상자의 협력에 대한 제약이 적고, 헬멧, 마스크 등으로 얼굴의 일부 영역이 가려진 상태에서도 획득이 가능한 장점이 있다[2, 3].

눈 주변 영역의 생체정보에 대한 초기 연구에서는 사람을 식별할 수 생체정보로서 눈 주변 영역에 대한 영상 획득 방법과 눈 영역의 설정 등에 대하여 탐색하였다. Park 외[3]는 눈 주변 영역에 홍채 또는 눈 영역에 대한 제외 여부, 눈썹 포함 여부 등 다양한 조건에서 개인 식별 성능을 비교하였는데, 눈 주변 영역에 눈 영역은 물론 눈썹도 포함시키는 것이 성능 향상에 도움이 되는 것으로 나타났다. 눈 주변 영역의 생체정보를 이용하여 사람을 식별할 때 성능에 영향을 미치는 요소를 탐색한 Padole 외[4]는 카메라와 사람사이의 거리를 1 m 간격으로 4 m부터 8 m까지 거리를 두고 획득한 영상들을 비교하여 7 m거리에서 획득한 영상이 가장 좋은 성능을 보였고, 포즈(pose) 변화에 따라 얼굴이 정면에서 벗어날 때 성능 저하가 발생함을 나타내었다. 그리고 눈 주변 영역을 지정하기 위한 기준점을 선정할 때 홍채의 중심점보다는 눈의 양쪽 코너점의 중간점이 보다 안정적인 것으로 나타났다. 홍채는 눈의 움직임에 따라 위치 변화가 크게 나타나기 때문이다. Smereka & Kumar [5]는 눈을 중심으로 눈썹, 눈 주위의 피부 등 사람을 식별하기 위하여 눈 주변 영역을 어느 정도까지 확장하는 것이 효과적인지 실험적으로 탐색하였다. 가시광선에서 획득한 영상의 경우는 눈 주변 영역을 넓게 확장하였을 때 가장 우수한 인식 성능을 나타냈으며, 근적외선에서 획득한 영상의 경우는 눈썹 일부만 포함되는 영역에서 가장 우수한 인식 성능을 나타내는 것으로 나타났다.

눈 주변 영역에 나타나는 특징을 추출하기 위하여 얼굴, 눈, 홍채 등 기존의 생체인식에서 사용되었던 특징 추출 방식을 비롯하여 다양한 특징 추출 방식에 대한 연구들이 진행되어 왔다. 국소 영역에 대한 특징을 효과적으로 추출하는 것으로 알려진 HOG(Histogram Of Gradient), LBP(Local Binary Pattern), SIFT(Shift Invariant Feature Transform) 특

징은 눈 주변 영역의 특징 추출에 널리 적용되어 왔다[1, 3, 4, 6]. Smereka & Kumar[5]는 두 영상을 5x5 개의 패치로 분할할 후 각 패치에 변형 모델을 적용하여 상관값이 가장 높게 나타나는 패치들로부터 얻은 상관 값을 이용하는 방식[7]과 수정된 SIFT 특징[8]을 이용한 방식을 적용하여 조명, 포즈 등의 변화에 강인한 인식을 시도하였다. Uzair 외[2]는 정규화된 영상에서 추출한 화소 값, 화소 값에 대한 주성분분석(PCA, Principal Component Analysis) 계수, LBP 특징, LBP 특징에 대한 PCA 계수라는 4개의 특징을 추출한 후 특징 레벨에서의 결합(feature score fusion)과 분류기 레벨에서의 결합(classifier score fusion)이라는 두 단계의 결합 방식을 제안하였다. 이 방식은 눈 주변 영역에서 추출한 다양한 특징을 결합하여 단일 특징이 갖는 한계를 극복하기 위한 것이다. 합성곱 특징을 제한된 볼츠만 머신(RBM, Restricted Boltzmann Machine)에 적용하여 CRBM(Convolutional RBM) 특징 추출 방식을 제안한 연구[9]와 공막(sclera)의 중심점을 기준으로 대칭 필터를 적용하여 여러 개의 링 영역에 나타나는 국소 정보를 추출하기 위한 SAFE(Symmetry Assessment by Feature Expansion) 특징 기술자를 제안한 연구[10]에서도 HOG, LBP, SIFT 특징과 결합하여 성능 향상을 도모하였다.

대용량의 영상 데이터인 ImageNet[11]을 기반으로 물체의 탐지 및 분류를 수행하는 대회에서 합성곱 신경망(CNN, Convolutional Neural Network)이 기존 방식에 비하여 매우 우수한 성능을 나타내면서[12], 생체인식 분야에서도 CNN을 적용한 연구들이 이루어지고 있다[13-15]. 그러나 아직까지는 눈 주변 영역에 대한 특징 추출에 CNN을 적용한 연구는 활성화되지 않은 상태이다. Kevin 외[16]는 ImageNet 영상으로 훈련된 기존의 다양한 CNN 모델을 이용하여 눈 주변 영역에 대한 특징을 추출한 후 유클리디안, 카이 제곱, 코사인 거리와 같은 단순한 특징거리 비교를 통해 효과적인 개인 식별이 가능함을 제시하였고, ResNet 모델이 AlexNet, GoogleNet에 비해 상대적으로 우수한 성능을 나타내었다. Zhao & Kumar [17]는 눈 주변 영역에서 의미 있는 정보를 내포하는 영역을 식별하기 위한 AttNet와 해당 영역에서 특징을 추출하는 FCN-Peri라는 두 개의 CNN 모델을 제안하였다.

본 연구에서는 상용 CNN 모델을 활용하여 소량의 학습 데이터로 눈 주변 영역의 특징을 추출하고, 이러한 특징을 이용하여 개인을 식별할 수 있는 방법을 제안하고자 한다. CNN의 깊이와 너비가 늘어날수록 인식 성능을 향상시킬 수 있을 것으로 기대되나 신경망의 깊이와 너비가 늘어날수록 학습을 위한 파라미터의 개수가 늘어나게 되고, 대용량의 학습 데이터를 확보하여야 한다. 그러나 눈 주변 영역에 대한 생체 데이터를 다량으로 획득하는 것은 현실적으로 해소하기 어려운 문제이다. 본 연구에서는 Kevin 외 [16] 연구에서 눈 주변 영역의 특징을 추출하는데 효과적인 것으로 보고된 ResNet 모델을 이용하되, 특징 추출을 위한 복잡한 심층 구조의 CNN 모델을 학습시키는 대신에 추출된 특징을 분류하기 위한 단순하고 얇은 구조의 다층퍼셉트론(MLP, Multi-layer Perception)을 학습시키는 방식을 사용하였다. 또한 소량의 학습 데이터로 MLP를 효과적으로 학습시키기 위하여 추출한 합성곱 특징을 PCA를 통해 차원을 축소시켜 MLP에 입력되도록 하였다. 이를 통해 상용 CNN 모델을 활용하여 소량의 학습 데이터로 기존의 전통적인 특징 추출 방법을 이용한 방식보다 우수한 개인 식별 성능을 얻을 수 있었다. 제안하는 방식은 생체 데이터와 같이 다량의 학습 데이터 확보가 어려운 응용에 복잡하고 깊은 CNN 모델을 적용할 수 있는 장점이 있다. 또한 상용 CNN 모델을 이용하기 때문에 인식 대상에 따라 적합한 CNN 모델을 선정하여 손쉽게 적용분야를 확장할 수 있는 장점이 있다. 본 논문의 구성은 다음과 같다. 2장에서는 ResNet 모델의 구조와 특징을 살펴보고, 3장에서는 ResNet 모델을 활용하여 눈 주변 영역의 특징을 추출하되, MLP 분류기를 이용하여 소량의 학습 데이터로 개인을 효과적으로 식별할 수 있는 방법에 대하여 설명한다. 4장에서는 공개 데이터베이스를 이용한 실험 결과를 제시하고, 기존 연구 결과와의 비교를 통해 제안된 방법의 타당성을 제시한다. 마지막으로 5장에서는 본 연구의 결론 및 향후 과제에 대하여 논의한다.

2. ResNet 모델

전통적 방식의 물체 분할 및 식별 방식에 비하여 AlexNet[12]이라는 심층 합성곱 신경망(DCNN) 모델을 이용한 방식이 주목할 만한 인식 성능을 나타냄

에 따라 깊은 신경망에 대한 연구가 촉발되었다. CNN의 기본 구조를 적층하여 저수준, 중간수준, 고수준의 특징 추출이 가능하고, CNN의 깊이와 너비가 늘어날수록 인식 성능을 향상시킬 수 있을 것으로 기대되기 때문이다. 그러나 신경망의 깊이와 너비가 늘어날수록 학습을 위한 파라미터의 개수와 연산 양이 늘어나는 문제를 해결하여야 한다. GoogleNet [18]은 큰 신경망을 희소(sparse)하게 연결한 구조인 Inception 모듈을 도입하여 학습시킬 파라미터의 양을 줄임으로써 신경망을 더욱 깊게 만들 수 있게 하였다.

ResNet[19]은 DCNN모델의 깊이가 늘어날수록 학습과정에서 오류가 증가함을 파악하고, 이러한 학습 오류를 최소화하기 위하여 바로가기 연결(short-cut connection)을 도입하였다. 최적의 학습 함수를 $H(x)$ 라고 할 때, Fig. 1에 제시된 $F(x)$ 는 $H(x)$ 에서 x 를 뺀 $F(x) := H(x) - x$ 로 나타낼 수 있다. 따라서 최적 함수를 $F(x)+x$ 로 나타낼 수 있다면 잔여 함수(residual function)를 통해 보다 쉽게 접근할 수 있다. $F(x)+x$ 는 Fig. 1에 제시된 바와 같이 바로가기 연결을 통한 전방향전달 신경망(feedforward neural network)로 실현될 수 있다.

이 방식은 기존 망의 구조를 크게 변경하지 않기 때문에 학습할 파라미터의 개수가 증가하지 않고, DCNN의 깊이가 더욱 깊어지더라도 학습 과정에서 역전파(back-propagation)되는 기울기(gradient)가 점차 사라져 학습이 잘 이루어지지 않는 기울기 소실 문제를 해결할 수 있는 장점이 있다. 이것은 전체 신경망에 걸쳐 가중치들을 최적화하는 것이 아니라, 2~3개의 레이어로 이루어진 잔여 블록(residual block)마다 부분적인 학습이 이루어지기 때문이다. ResNet50 모델은 Fig. 2와 같이 3개의 레이어마다 바로

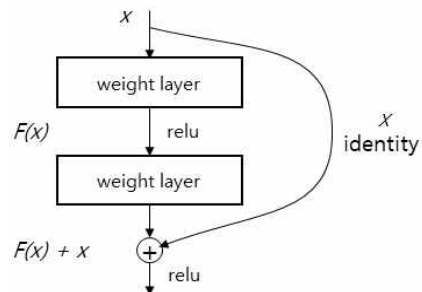


Fig. 1. Residual learning: a building block [19].

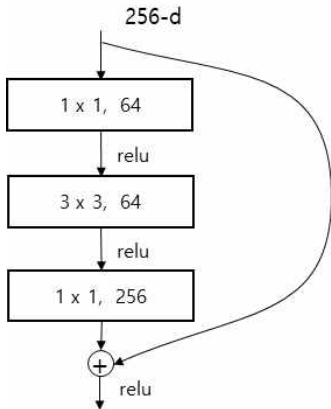


Fig. 2. Bottleneck building block for ResNet50 [19].

가기 연결이 되어 있으며, 각 잔여 블록은 1×1, 3×3, 1×1 합성곱 레이어로 구성되어 있다. 1×1 합성곱 레이어는 차원을 축소하고 나중에 복원함으로써 3×3 합성곱 레이어의 입력력 차원을 축소하여 병목(bottleneck)을 구성한다. Table 1은 ResNet50의 전체적인 구조를 나타낸 것으로, conv3_1, conv4_1, conv5_1에서도 스트라이드(stride) 2로 다운샘플링이 진행된다.

3. ResNet50 모델을 이용한 특징 추출 및 인식

ResNet50 모델은 50개의 합성곱 레이어를 갖는

Table 1. Architecture of ResNet50

Layer name	Output size	ResNet50
conv1	112×112	7×7, 64, stride 2
conv2_x	56×56	3×3 max pool, stride 2
		$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
conv4_x	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
conv5_x	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	avg. pool, 1000d FC

DCNN으로 눈 주변 영역의 특징을 추출하여 개인을 식별하기 위해서는 학습에 필요한 대용량의 생체 정보를 확보하여야 한다. 그러나 눈 주변 영역에 대한 생체 정보를 대량으로 획득하는 것은 현실적으로 불가능하다. ImageNet으로 사전 학습된 ResNet50 모델의 가중치를 목적 데이터인 눈 주변 영역 영상 데이터를 이용하여 미세조정(fine tuning)하는 방법이 있으나, 생체정보의 경우에는 개인 별로 획득 가능한 데이터의 양이 매우 제한적이기 때문에 이 방식을 적용하여 원하는 성능을 기대하기는 어렵다. 따라서 본 연구에서는 ImageNet 데이터로 사전 학습된 ResNet50 모델의 합성곱 특징을 이용함으로써 깊은 신경망을 학습시켜야 하는 문제를 해소하고, 그 대신에 합성곱 특징으로부터 개인을 효과적으로 식별하기 위하여 얇은 MLP를 학습시키는 방식을 사용하였다. 그리고 매우 제한된 양의 생체 데이터를 이용하여 MLP 분류기를 효과적으로 학습시키기 위해서 추출한 합성곱 특징을 PCA를 통해 차원을 축소시켜 MLP에 입력되도록 하였다. Fig. 3은 ResNet50 모델을 이용한 눈 주변 영역의 특징을 추출한 후 MLP를 이용하여 개인을 식별하는 과정을 도식화 한 것이다.

방대한 ImageNet 데이터를 통해 학습된 ResNet50 모델은 1,000개의 물체를 식별하도록 설계되고 학습되었으므로, 이 모델을 눈 주변 영역의 정보를 이용한 생체인식에 적용하기 위해서는 입력 영상의 크기 정규화를 비롯하여 주어진 도메인에 적합한 특징을 추출하기 방안이 요구된다. 눈 주변 영역의 크기는 카메라와 사람간의 거리에 따라 달리 나타나므로 224×224로 크기 정규화를 수행하였다. 크기 정규화 후 RGB 컬러영상, 명도 영상, Gabor 필터를 적용한 영상을 ResNet50 모델의 입력으로 사용하고, 최상단 레이어만 현재 도메인에 맞게 수정한 후 가중치를 미세 조정하는 학습을 통해 식별 성능을 살펴보았다. 그 결과 명도 영상이 눈 주변 영역 특징을 추출하는 데 적합한 것으로 나타났으며, 일반적으로 특징 추출 과정에서 성능 향상에 기여하는 것으로 보고되었던 Gabor 필터를 적용한 영상은 식별 성능이 가장 낮게 나타나 원본 영상에 인위적으로 특징을 부각시키는 방법은 DCNN 모델에는 역효과를 내는 것으로 판단된다.

일반적으로 DCNN의 앞쪽에 나타나는 합성곱 특징은 색상이나 에지와 같은 낮은 차원의 특징을 식별

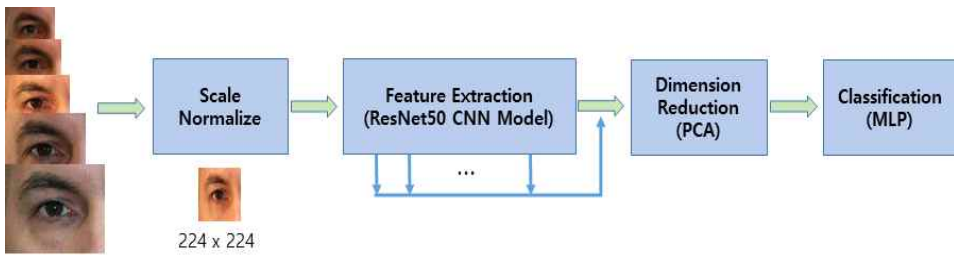


Fig. 3. Overview of feature extraction and classification.

하고, 뒤쪽에서는 이런 낮은 차원의 특징을 결합하여 보다 고차원적인 물체의 형태적 특징을 추출하는 것으로 알려져 있다. 눈과 눈썹의 위상적 관계는 모든 사람이 동일할 뿐만 아니라 눈과 눈썹의 모양도 전체적으로는 크게 다르지 않다. 다만 사람마다 눈썹의 형태, 쌍꺼풀, 눈 부위의 골격 등에 따라 부분적으로 그 형태의 차이가 나타날 뿐이다. 따라서 ResNet50 모델의 최종 레이어에 비하여 중간 및 앞쪽 레이어에서 추출한 합성곱 특징이 눈 주변 영역의 특징을 추출하는데 효과적일 수 있다. 실험을 통해 눈 주변 영역의 특징을 추출하는데 적합한 ResNet50 모델의 합성곱 특징을 탐색한 결과, 중간 부분에 위치한 합성곱 특징이 눈 주변 영역에 의한 개인 식별에 적합한 것으로 파악되었다.

합성곱 레이어에서 추출한 특징이 동일한 클래스에서 유사하게 나타나고, 상이한 클래스에서 차이가 크게 나타날수록 개인 식별에 대한 성능 향상을 기대할 수 있다. 그러나 영상 획득 시 거리나 포즈의 변화에 따라 눈 주변 영역에 대한 특징에 변이가 발생하므로 안정적인 특징을 추출하기 위하여 합성곱 레이어의 출력을 7×7 크기의 평균 풀링을 적용하였다. 또한 풀링 연산은 특징의 차원을 축소시키는 역할도 수행하게 된다. 매우 제한된 양의 생체 데이터를 이용하여 MLP 분류기를 효과적으로 학습시키기 위해서 추출한 합성곱 특징을 PCA를 통해 차원을 축소시켜 MLP에 입력되도록 하였다. PCA는 무작위 특이값 분해(SVD, Singular Value Decomposition)를 적용하였고, MLP 분류기의 활성화 함수는 relu, 가중치 최적화를 위해서는 adam을 사용하였다. 그리고 눈 영역 특징 분류에 적합한 MLP 분류기의 은닉층 개수와 노드의 개수는 실험을 통해 설정하였다. Fig. 4는 본 연구에서 사용된 은닉층이 1개인 MLP 구조를 제시한 것이다.

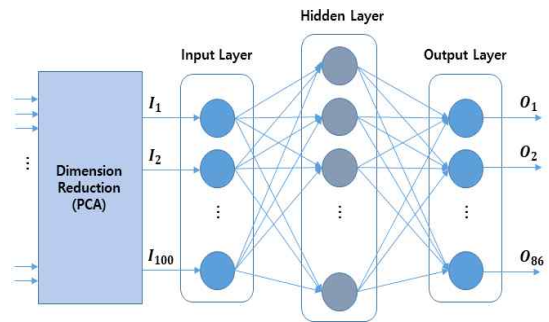


Fig. 4. Multi-layer perceptron with one hidden layer.

4. 실험 결과 및 분석

4.1 실험 데이터

사전 학습된 ResNet50 모델을 이용하여 눈 주변 영역의 특징을 추출하고 MLP 분류기를 학습시켜 개인을 식별하는 제안된 방법의 타당성을 확인하고, 그 결과를 기존 연구들과 비교하기 위하여 공개된 데이터베이스인 UBIPr 데이터세트를 사용하였다. UBIPr 데이터세트는 두 세션에 걸쳐 데이터가 수집되었으며, 데이터를 수집할 때 조명의 변화는 물론 카메라와 피사체 간의 거리에 변화를 주었다. Fig. 5에 제시된 바와 같이 카메라와 피사체 간의 거리를 4 m부터 8 m까지 1 m 간격으로 5가지 거리 변화를 주어 촬영하였다. 본 실험에서는 두 개의 세션에 걸친 영상 획득에 참여한 86명의 데이터, 총 1,718개의 눈 주변 영역



Fig. 5. Images from the UBIPr database: captured at different distances

영상을 이용하여 실험을 수행하였다. 2개의 세션, 5개의 거리에서 영상을 획득하였기 때문에 1인 당 획득한 영상은 왼쪽 10개, 오른쪽 10개이다. 86명 중 1명의 데이터는 왼쪽, 오른쪽 각각 8개 이다.

4.2 인식 실험 및 결과

인식 실험은 두 단계로 나누어 실시하였다. 첫 단계는 ResNet50 모델의 CNN 레이어별 출력 특징, PCA에 의한 차원 축소, MLP 분류기의 은닉층의 개수와 노드의 개수 등에 따른 성능 변화를 알아보기 위한 실험으로, 첫 번째 세션에서 획득한 데이터 중 왼쪽 눈 영상 데이터를 이용하였다. MLP 분류기의 학습을 위하여 60%의 데이터를 사용하고, 평가를 위해 나머지 40%의 데이터를 사용하였다. 둘째 단계의 실험은 첫 단계의 실험 결과에 따라 인식 대상 도메인에 따른 파라미터를 설정한 후 첫 번째 세션에서 획득한 데이터를 학습 데이터로 사용하고, 한 번도 사용하지 않았던 두 번째 세션에서 획득한 데이터를 평가 데이터로 사용하였다. 이러한 실험 설계는 눈 주변 영역의 생체 정보를 활용하여 개인을 확인하는 실제 응용을 고려한 것이다.

생체인식 시스템에서 개인 인증은 인증을 요청한 사람으로부터 추출한 특징을 인증시스템에 등록된 특정인의 특징과 비교하여, 미지의 개인이 특정인과 동일인인지 아닌지를 판정한다. 인증을 요청한 사람이 등록된 특정인과 동일인이 아님에도 불구하고 동일인으로 판단하는 비율을 오수락률(FAR, False Acceptance Rate)이라 하고, 동일인임에도 불구하고 동일인이 아니라고 판단하는 비율을 오거부율(FRR, False Reject Rate)라고 한다. FAR를 줄이기 위해서는 동일인임을 판단하는 기준을 엄격하게 적용해야 하는데, 이렇게 하면 FRR이 늘어나게 된다. 반대로 FRR을 줄이기 위해서 동일인임을 판단하는 기준을 느슨하게 적용하면 FAR이 늘어나게 된다. 이러한 이유로 FAR과 FRR이 일치하는 지점에서의 오류율을 나타내는 동일오류율(EER, Equal Error Rate)이 생체인식 시스템의 성능을 측정하는 대표적인 척도로 사용된다. 그러므로 본 연구에서도 EER을 사용하여 제안된 방식의 성능을 측정하였다.

ResNet50 모델은 총 177개의 레이어로 구성되어 있다. 바로가기로 연결된 블록을 중심으로 살펴보면 50개 블록으로 나누어 볼 수 있고, 블록의 CNN 출력

의 크기로 살펴보면 5개 그룹으로 구분할 수 있다. 어떤 레이어의 CNN 특징이 눈 주변 영역을 식별하는데 보다 효과적인지를 파악하기 위하여 각 그룹별 최종 레이어의 CNN 특징을 추출하여 실험에 사용하였다. CNN 특징은 7×7 크기의 평균 풀링을 거친 후 PCA를 통해 차원을 축소하여 MLP 분류기의 입력력으로 보내진다. MLP 분류기의 활성화 함수는 relu, 가중치 최적화를 위해서는 adam을 사용하였으며 최대 반복 횟수는 200으로 설정하였다. Fig. 6은 100개의 노드로 구성된 1개의 은닉 층을 갖는 단순한 구조의 MLP에 입력되는 특징의 차원을 PCA의 의해 25, 50, 100, 200, 400으로 변화시켰을 때의 EER의 변화를 나타낸 것이다. acti_1은 Table 1에 제시된 conv1에 해당하는 레이어고, acti_10, acti_22, acti_40, acti_49는 각각 conv2_x, conv3_x, conv4_x, conv5_x의 최종 레이어를 나타낸다. ResNet50 모델의 중간 부분에서 추출한 특징의 식별력이 좋은 것으로 나타났으며, 입력 특징의 차원이 50, 100 일 때 특징을 추출한 CNN 레이어에 무관하게 상대적으로 좋은 성능을 나타냈다.

MLP에 입력되는 특징의 차원을 고정시킨 후 MLP의 은닉 층의 개수를 2개, 3개로 늘려가면서 성능을 측정한 결과 오히려 성능이 저하되는 것으로 나타났다. 이러한 현상은 은닉 층의 노드 개수를 줄이거나 늘려도 유사한 결과를 나타내었다. 이것은 은닉 층이 늘어나면서 학습이 효과적으로 이루어지지 않았기 때문으로 판단된다. Fig. 7은 MLP의 은닉 층은 1개로 유지한 상태에서 은닉 층 노드의 개수를 변화시켜가면서 성능을 측정한 결과이다. 노드의 개수를 늘렸을 때 인식 성능이 향상되는 것으로 나타났으며 400개에서 800개로 확장했을 때 acti_22레이어

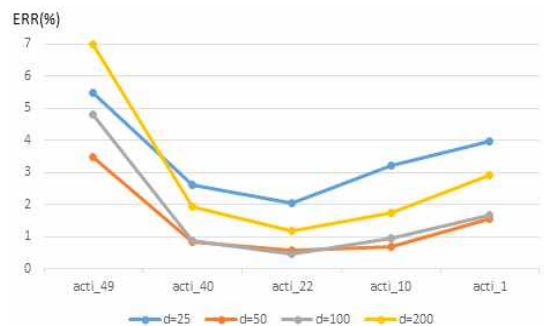


Fig. 6. Performance variation with different dimensions.

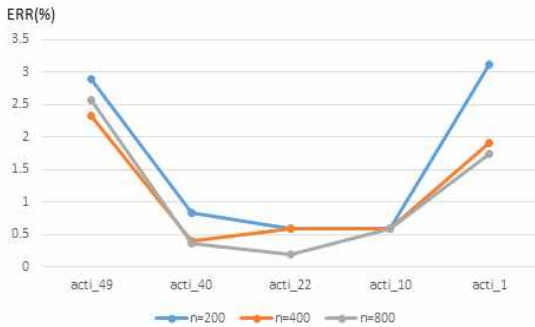


Fig. 7. Performance variation with different number of nodes.

를 제외하고는 성능 향상이 크게 나타나지는 않았다.

Table 2는 생체인식 시스템이 실제로 운용되는 환경과 유사하게 첫 번째 세션에서 수집한 데이터를 학습데이터로 하여 MLP 분류기를 훈련시킨 후 한번도 본적이 없는 두 번째 세션에서 수집한 데이터로 인증 테스트를 진행한 결과이다. ResNet50 모델에서 conv3_x 그룹의 마지막 레이어인 acti_22에서 추출한 특징을 PCA를 통해 d -차원으로 축소하고 은닉층에 n 개의 노드를 갖는 MLP 분류기를 학습시켜 얻은 결과이다. ResNet50 모델을 통해 추출한 합성곱 특징을 PCA를 통해 50차원으로 축소시켜 MLP 분류기를 학습시켰을 때, 왼쪽 눈 영역에 비하여 오른쪽 눈 영역에서 성능 저하가 나타났으나 100차원으로 특징을 확대하였을 때는 이러한 현상이 줄어들었다. 이것은 PCA의 의해 특징을 50차원으로 축소할 경우 유의미한 식별 정보가 상실됨을 의미한다.

4.3 인식 성능 비교

눈 주변 영역 인식을 위한 제안된 방법의 효과성을 비교하기 위하여 동일한 실험 데이터를 이용한 최근의 연구 결과들을 정리하여 Table 3에 제시하였다. 눈 주변 영역을 이용하여 개인을 식별할 때 특징 추출에 적합한 눈 주변 영역의 범위를 실험적으로

Table 2. Verification result (EER) of periocular region

PCA	MLP	Left	Right
d=50	n=400	2.33%	4.42%
	n=800	2.31%	4.41%
d=100	n=400	2.56%	2.56%
	n=800	2.49%	2.72%

Table 3. Performance comparison with different researches

Dataset		Left	Right
Research			
Semereka & Kumar[5]		7.62%	6.43%
Fernando et.al.[10]		7.9%	
Kevin et.al.[16]		5.1%	
Zhao & Jumar[17]		2.26%	
Proposed method	d=50	2.31%	4.41%
		2.79%	
	d=100	2.49%	2.72%
		2.10%	

탐색한 Smereka & Kumar[5]의 연구에서는 왼쪽 눈은 왼쪽 눈, 오른쪽 눈은 오른쪽 눈과 비교하였다는 것 외에 실험 방식에 대한 구체적인 설명을 제시하지 않았다. 이것은 동일한 조건에서 눈 주변 영역으로 설정한 4개 영역별 상대적인 비교에 초점을 두었기 때문인 것으로 생각된다. Fernando 외[10]는 전통적인 특징 추출 방식인 HOG, LBP, SIFT 방식에 대칭 필터를 적용한 SAFE 기술자를 결합하였고, Kevin 외[16]는 ImageNet 데이터베이스로 학습된 CNN을 이용하여 특징을 추출한 후 단순한 유사도 비교 방식으로 인식을 수행하였다. [10]과 [16]은 본 연구의 실험 방식과 동일하게 첫 번째 세션에서 획득한 데이터로 학습을 수행하고, 두 번째 세션에서 획득한 데이터로 인증 실험을 수행하였다. 다만 86명에게서 획득한 왼쪽 눈과 오른쪽 눈을 서로 다른 사람에게서 획득한 생체 정보로 간주하여 분류할 클래스를 172개로 늘려서 인증 실험을 진행하였다. 눈 주변 영역에서 의미 있는 영역을 식별하기 위한 AttNet와 해당 영역에서 특징을 추출하는 FCN-Peri라는 두 개의 CNN 모델을 제안한 Zhao & Jumar[17]의 연구는 왼쪽 눈만을 이용하였고, 두 세션에 존재하는 동일인 이외 다른 사람들도 포함시켜 인증 실험을 진행하였다.

본 연구에서는 첫 번째 세션에서 획득한 왼쪽 눈과 오른쪽 눈 데이터에 대해 각각 대응되는 두 번째 세션의 데이터를 이용한 인증 실험은 물론이고, 동일한 조건에서의 객관적인 성능 비교를 위하여 [10]과 [16]에서 수행한 172개 클래스로 분류하는 방식으로 인증 실험을 진행하여 그 결과를 Table 3에 제시하였다. 제시된 결과는 PCA를 통해 특징을 50차원과

100차원으로 축소한 후 800개의 노드를 갖는 MLP 분류기를 학습시켜 얻은 결과이다. 기존 연구 결과들과 인식 성능을 비교하였을 때, 기존의 전통적인 특징 추출 방법을 이용한 방식[5, 10]은 물론 CNN을 이용한 Kevin 외[16]의 방식보다 우수한 개인 식별 성능을 나타내었다. 왼쪽 눈 영상에 대하여 Zhao & Jumar[17] 방식 보다는 EER이 약간 높게 나타났으나, 최고의 인식 성능을 보인 기존 연구에 버금가는 성능을 나타냈다. 제안하는 방식은 상용 CNN 모델을 이용하기 때문에 인식 대상에 따라 적합한 CNN 모델을 선정하여 손쉽게 적용분야를 확장할 수 있는 장점이 있다.

5. 결 론

눈 주변 영역의 생체 정보는 기존의 지문, 홍채 등과는 달리 근접하지 않아도 생체 정보의 획득이 가능하며, 인식 대상자의 협력에 대한 제약이 적고, 헬멧, 마스크 등으로 얼굴의 일부 영역이 가려진 상태에서도 획득이 가능한 장점이 있다. 기존의 연구들은 눈 주변 영역에 나타나는 특징을 효과적으로 추출하기 위하여 HOG, LBP, SIFT 등의 전통적인 특징 추출 방법을 능가할 수 있는 특징을 발굴하고, 이들 특징들을 효과적으로 결합할 수 있는 방법들에 대하여 집중하여 왔다. 최근 들어 CNN을 이용한 특징 추출 방법이 여러 방면에서 우수한 성과를 나타냄에 따라 CNN을 이용하여 생체 인식을 수행하는 연구가 이루어지고 있다. 그러나 CNN을 적용하여 눈 주변 영역의 생체 정보를 식별하는 연구는 아직까지 저조한 상태이다. 이것은 심층신경망을 훈련시키는데 필요한 충분한 양의 데이터를 구하기 어려운 현실적 측면이 있기 때문인 것으로 판단된다. 본 연구에서는 이러한 문제를 해결하기 위하여 1,000개의 클래스, 약 3백만 개의 영상들로 학습된 ResNet50 모델을 이용하여 눈 주변 영역의 특징을 추출하고 MLP 분류기를 이용하여 개인 인증을 효과적으로 수행하는 방법을 제안하였다. 공개된 UBIPr 데이터셋을 이용하여 실험한 결과 기존 연구 결과들과 비교하였을 때 매우 효과적임을 확인할 수 있었다. 향후 연구에서는 기존의 전통적인 특징 추출 방식으로 추출한 특징과 CNN을 이용하여 추출한 특징을 결합하여 개인 인증 성능을 향상시킬 수 있는 방법에 대하여 추가적인 연구를 진행할 계획이다.

REFERENCE

- [1] F.P. Mahdi, M. Habib, A.R. Ahad, S. Mckeever, A.S.M. Moslehuddin, P. Vasant, "Face Recognition-based Real-time System for Surveillance," *Intelligent Decision Techniques*, Vol. 11, No. 1, pp. 79-92, 2017.
- [2] M. Uzair, A. Mahmood, A. Mian, and C. McDonald, "Periocular Region-based Person Identification in the Visible, Intrared and Hyperspectral Imagery," *Nurocomputing*, Vol. 149, pp. 854-867, 2015.
- [3] U. Park, R.R. Jillela, A. Ross, and A.K. Jain, "Periocular Biometrics in the Visible Spectrum," *IEEE Transaction on Information Forensics and Security*, Vol. 6, No. 1, pp. 96-106, 2011.
- [4] C.N. Padole and H. Proenca, "Periocular Recognition: Analysis of Performance Degradation Factors," *Proceedings of International Conference on Biometrics*, pp. 439-445, 2012.
- [5] J.M. Smereka and B.V. Kumar, "What is a 'Good' Periocular Region for Recognition?" *Proceedings of the IEEE International Workshop on Computer Vision and Pattern Recognition*, pp. 117-124, 2013.
- [6] K.K. Kumar and M. Pavani, "LBP Based Biometric Identification Using the Periocular Region," *Proceedings of IEEE Annual Information Technology, Electronics and Mobile Communication Conference*, pp. 204-209, 2017.
- [7] V.N. Boddeti, J.M. Smereka, and B.V. Kumar, "A Comparative Evaluation of Iris and Ocular Recognition Methods on Challenging Ocular Images," *Proceedings of International Conference on Biometrics*, pp. 1-8, 2011.
- [8] A. Ross, R. Jillela, J.M. Smereka, V.N. Boddeti, B.V. Kumar, R. Barnard, et al., "Matching Highly Non-ideal Ocular Images: An Information Fusion Approach," *Proceedings of International Conference on Biometrics*, pp. 446-453, 2012.
- [9] L. Nie, A. Kumar, and S. Zhan, "Periocular

- Recognition Using Unsupervised Convolutional RBM Feature Learning,” *Proceedings of International Conference on Pattern Recognition*, pp. 399–404, 2014.
- [10] A.F. Fernando, M. Anna, and B. Josef, “Compact Multi-scale Periocular Recognition Using SAFE Feature,” *Proceedings of International Conference on Pattern Recognition*, pp. 1455–1460, 2016.
- [11] D. Jia, D. Wei, S. Richard, L.J. Li, K. Li, F.F. Li., “ImageNet: A Large-scale Hierarchical Image Database,” *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 248–255, 2009.
- [12] A. Krizhevsky, I. Sutskever, and G.E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Proceedings of International Conference on Neural Information Processing System*, Vol. 25, No. 2, pp. 1097–1105, 2012.
- [13] M. Coskun, A. Ucar, O. Yidirim, Y. Demir, “Face Recognition Based on Convolutional Neural Network,” *Proceedings of International Conference on Modern Electrical and Energy Systems*, pp. 376–379, 2017.
- [14] M. Kim, “Contactless Palmprint Identification Using the Pretrained VGGNet Model,” *Journal of Korea Multimedia Society*, Vol. 21, No. 12, pp. 1439–1447, 2018.
- [15] K. Nguyen, C. Fookes, A. Ross, and S. Sridharan, “Iris Recognition With Off-the-Shelf CNN Features: A Deep Learning Perspective,” *IEEE Access*, Vol. 6, pp. 18848–18855, 2017.
- [16] H.D. Kevin, A.F. Fernando, and B. Josef, “Periocular Recognition Using CNN Feature Off-the-Shelf,” *Proceedings of International Conference on Biometrics Special Interest Group*, pp. 1–5, 2018.
- [17] Z. Zhao and A. Jumar, “Improving Periocular Recognition by Explicit Attention to Critical Regions in Deep Neural Network,” *IEEE Transaction on Information Forensics and Security*, Vol. 13, No. 12, pp. 2937–2952, 2018.
- [18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., “Going Deeper with Convolutions,” *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.
- [19] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.



김민기

1989년 2월 중앙대학교 전자계산학과 이학사
 1994년 8월 중앙대학교 전자계산학과 이학석사
 1998년 2월 중앙대학교 컴퓨터공학과 공학박사

1998년~1999년 KRIC 선임연구원
 1999년~2000년 KERIS 책임연구원, 연구위원
 2000년~2010년 경상대학교 컴퓨터교육과 부교수
 2011년~현재 경상대학교 컴퓨터과학과 교수
 2013년~2014년 미국 University of Notre Dame 방문학자
 2014년~현재 경상대학교 공학연구원(ERI) 자동화·컴퓨터연구센터 책임연구원
 관심분야 : 딥 러닝, 생체인식, 패턴인식, 영상처리