

CNN을 적용한 조명변화에 강인한 얼굴인식 연구

김연호* · 박성욱* · 김도연**

Research on Robust Face Recognition against Lighting Variation using CNN

Yeon-Ho Kim* · Sung-Wook Park* · Do-Yeon Kim**

요약

얼굴인식 기술은 지난 수십 년간 연구되어온 분야로서 보안, 엔터테인먼트, 모바일 서비스 등 다양한 영역에서 활용되고 있다. 얼굴인식 기술이 가진 주된 문제점은 밝기, 조명각도, 영상 회전등의 환경적 변화 요소에 따라 인식률이 현저하게 감소된다는 것이다. 따라서 본 논문에서는 최근 많은 계산량을 처리할 수 있는 컴퓨터 하드웨어와 알고리즘의 발전으로 재조명 받고 있는 CNN을 이용해 조명변화에 강인한 얼굴인식 방법을 제안하였다. 이후 성능검증을 위해 기존의 얼굴인식 알고리즘인 PCA, LBP, DCT와 결과 비교를 진행하였으며, 각각 9.82%, 11.6%, 4.54%의 성능 향상을 보였다. 또한 기존 신경망을 적용한 얼굴인식 연구결과 비교에서도 5.24%의 성능 향상을 기록하여 최종 인식률 99.25%를 달성하는 결과를 보였다.

ABSTRACT

Face recognition technology has been studied for decades and is being used in various areas such as security, entertainment, and mobile services. The main problem with face recognition technology is that the recognition rate is significantly reduced depending on the environmental factors such as brightness, illumination angle, and image rotation. Therefore, in this paper, we propose a robust face recognition against lighting variation using CNN which has been recently re-evaluated with the development of computer hardware and algorithms capable of processing a large amount of computation. For performance verification, PCA, LBP, and DCT algorithms were compared with the conventional face recognition algorithms. The recognition was improved by 9.82%, 11.6%, and 4.54%, respectively. Also, the recognition improvement of 5.24% was recorded in the comparison of the face recognition research result using the existing neural network, and the final recognition rate was 99.25%.

키워드

Face Recognition, Deep Learning, Yale Face Database B, Convolutional Neural Network
얼굴 인식, 딥러닝, 예일 페이스 데이터 베이스 B, 컨볼루션 신경망

1. 서론

얼굴인식 알고리즘은 지난 수십 년간 인식성능 향

상을 위해 연구되어왔다. 이러한 노력을 통해 얼굴인식 기술은 많은 발전을 이루었으며, 보안, 엔터테인먼트, 모바일 서비스 등 다양한 분야에서 활용되고 있

* 순천대학교(kimyh7102@sunchon.ac.kr, park7231654@naver.com)

** 교신저자 : 순천대학교 컴퓨터공학과

• 접수일 : 2017. 02. 02

• 수정완료일 : 2017. 04. 13

• 게재확정일 : 2017. 04. 24

• Received : Feb. 02, 2017, Revised : Apr. 13, 2017, Accepted : Apr. 24, 2017

• Corresponding Author : Do-Yeon Kim

Dept. of Computer Engineering, Suncheon National University,

Email : dykim@sunchon.ac.kr

다. 하지만 얼굴인식은 밝기, 조명각도, 원본 영상의 회전등의 가변적 요소가 포함된 경우 인식률이 현저하게 저하되는 문제가 발생하게 된다. 이에 따라 가변적 요소에 강건한 알고리즘을 개발하기 위한 지속적인 노력이 진행되고 있다.

최근 기계학습 알고리즘의 한 종류인 딥러닝이 많은 계산량을 처리할 수 있는 하드웨어와 알고리즘의 발전으로 다시 재조명 받으며 특히 컴퓨터비전, 영상처리 등의 분야에서 눈부신 성능향상을 보이고 있다. 딥러닝은 다수의 계층구조로 구성된 모델의 사용과 여러 추상화를 통해 스스로 데이터의 특징 패턴을 학습[1]하고 높은 인식성능을 보여주고 있으며 환경변화에 따른 가변적 요소에도 강인한 인식성능을 보이고 있다. ILSVRC[2] 대회에서는 기존의 특징추출 알고리즘인 SIFT[3], SURF[4]등과 겨루어 고수준의 특징추출을 통한 높은 인식성능을 보이며 딥러닝 알고리즘의 우수성을 검증했다.

본 논문에서는 딥러닝을 이용해 조명각도에 따른 얼굴 인식을 저하 문제를 해결하기 위한 연구를 진행하였다. 얼굴인식을 위한 학습모델 구동을 위해 Caffe[5] 프레임워크를 사용했으며, 학습모델은 CNN 모델중 하나인 VGG-16[6] 네트워크 구조를 근간으로 네트워크 제작 및 세부적인 파라미터 조정을 통한 학습 망을 구성하였고, 학습을 위한 데이터베이스는 다양한 조명 밝기의 얼굴영상으로 구성된 Yale Face Database B[7]을 사용하여 가변적 요소에 강건한 알고리즘을 제안한다. 논문 2장에서는 얼굴인식 관련 연구를 소개하였으며, 3장에서는 CNN의 기본 구조의 설명과 제안하는 CNN 학습모델 구성을 기술하였다. 4장에서는 학습데이터베이스 구성, 실험결과 및 성능평가를 기술하였으며, 마지막 5장에서는 결론 및 향후과제를 기술하였다.

II. 관련연구

주성분 분석(principal component analysis, PCA)[8]은 D차원의 특징 벡터 정보 손실을 최소로 유지하고 차원을 축소시킬 목적으로 사용된다. 차원 축소를 위해 입력 공간에서 전체 성분들의 주축을 찾은 뒤 해당 축에 대해 투영시켜 균일 하게 분포된 데이터 값을 얻어낸다. 계산 방법은 전체 데이터에서 그

데이터의 평균을 구해 뺄셈 연산을 적용한 후, 공분산 행렬을 만들어 이를 이용해 고유 값과 고유 벡터를 구한다. 다음으로 전체 데이터를 고유 벡터에 투영해 나온 결과 값을 얼굴 인식을 위한 특징벡터[9]로 사용한다. PCA는 입력되는 데이터의 종류에 구애 받지 않는다는 장점이 있지만 얼굴의 표정, 조명, 대칭 등의 부분적인 요소의 변형에 대해서 대처가 미흡하다는 문제점이 있다.

지역이진패턴(local Binary Pattern, LBP)[10]은 입력 영상에서 256차원의 특징 벡터를 생성하는 텍스처 기반의 알고리즘이다. LBP는 영상의 모든 픽셀에 3X3 커널을 통해 현재 화소 값과 여덟 개 이웃의 값을 비교하여 0과 1로 변환한다. 이후 LBP를 얼굴인식에 활용한 예[11]를 보면, 얼굴 영역을 일정한 크기의 화소로 분할한 후 각 화소마다 LBP에 대한 히스토그램을 구하고, 구해진 히스토그램들을 일렬로 연결한 벡터를 최종 특징으로 사용한다. LBP는 이웃 화소간의 상대적인 명암 크기를 사용하므로 조명과 회전에 강인한 좋은 특성을 지니지만 명암이 비교적 균일한 부근에서 불안정하며 값을 출력할 때 크기 변화에 적절히 대응하지 못한다는 단점이 있다.

이산 코사인 변환(discrete cosine transform, DCT)[12]은 LBP와 같이 조명 변화에 불변인 특성을 지니지만, 주파수 정보를 이용하여 특징을 추출한다는 점이 다르다. 영상 기반 모델링 알고리즘[13]은 얼굴 특징 점의 고유한 특성인 반사율이 고려되지 않아 얼굴 인식을 수행 할 때, 조명의 밝기가 점차 어두워지면 급격한 성능저하를 보인다. DCT는 자연적인 영상을 주파수 성분별로 분해하여 저주파수 형태의 신호와 고주파수 형태의 신호로 나누어 얼굴 성분을 계수로 출력해 특징 점으로 사용한다. 이러한 DCT는 얼굴 인식에서 영상 기반 모델링 알고리즘과 비교하였을 때 조명 변화에 강인하며 인식률에서도 개선된 성능을 보이지만, 가변적 요소에 강인한 인식률을 보이지 못한다는 한계가 있다.

III. 제안하는 CNN기반 학습모델 구성

3.1 CNN의 기본 구조

컨볼루션 신경망(Convolutional neural network)은

다층 인공신경망의 일종으로 이미지와 같이 각 데이터 차원간 기하학적 연관성을 가지는 데이터를 효과적으로 인식하기 위해 제안되었다[14].

CNN의 기본 구조는 보통 convolution layer, pooling layer, fully connected layer와 같은 3가지 계층을 깊게 쌓아 신경망의 성능을 높이게 된다. convolution layer는 입력되는 전체 영상에서 convolution 연산을 거쳐 feature map 추출을 진행한다. pooling layer는 convolution Layer에서 추출된 특징을 바탕으로 sub-sampling을 사용으로 차원 축소와 입력 공간의 추상화를 통해 약한 특징은 무시되며 보다 강한 특징을 추출한다. fully connected layer는 convolution layer, pooling layer의 반복을 통해 추출된 특징들을 이용한 객체의 카테고리별 분류 목적으로 사용된다. 이후 마지막 레이어부터 초기 레이어까지 역전파 (back-propagation) 알고리즘의 사용으로 오차를 최소화하기 위해 가중치를 찾는 학습 최적화를 진행하며, 지속적인 반복학습을 통해 점차적으로 강한 특징 맵 추출과 높은 정확도의 분류율을 도출한다.

3.2 학습 네트워크 구성

본 장에서는 VGG-16 네트워크를 근간으로 추가된 학습 레이어를 통해 학습 데이터셋에 따라 달라지는 최적의 네트워크 구조를 38명에 맞춰 제안한다.

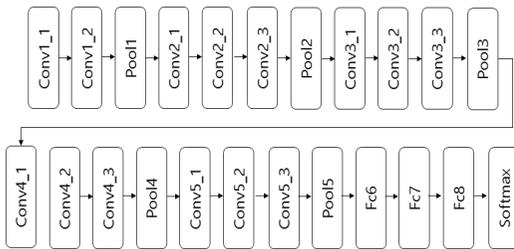


그림 1. 제안된 네트워크 구조
Fig. 1 Proposed network architecture

네트워크 최적화를 위해 기존의 13개의 convolution layer에서 1개의 convolution layer가 추가된 14개의 convolution layer를 사용하였으며, 이후 5개의 pooling layer와 3개의 fully connected layer의 사용으로 학습 네트워크를 그림1과 같이 구성하였다.

Conv1_1 convolution layer 단계에서는 입력되는

이미지로부터 kernel_size : 3x3, pad : 1의 값을 적용시킨 convolution 연산을 거쳐 feature map을 추출한다. 이어서 활성화함수(ReLU)의 사용으로 신경망의 출력을 곡선의 형태로 바꾸어 연속적인 미분을 통해 성능향상 역할을 수행한다. Conv1_2 단계에서는 Conv1_1과 동일한 파라미터 값을 적용시킨 컨볼루션 연산 수행 후, 다시 활성화함수를 사용하였다. 이후 Pool1 풀링 레이어 단계에서는 sub-sampling을 통해 입력 이미지 해상도를 절반으로 줄이며, convolution layer에서 추출된 특징들로부터 약한 특징은 무시하고 강한 특징들을 추출하는 max_pooling 작업을 진행하게 된다.

Conv2_1 단계에서는 Pool1 단계에서 추출된 특징들을 다시 3x3 크기의 커널과 패딩값을 1로 적용시킨 convolution 연산을 거쳐 feature map을 추출과 활성화함수의 사용을 Conv2_2, Conv2_3에서도 반복 진행한다. 다음 Pool2 단계에서는 Pool1과 같이 convolution layer에서 추출된 feature map을 통해 sub-sampling을 거쳐 강한 특징을 추출하게 된다. 이후 반복적인 convolution layer와 pooling layer의 연산을 거쳐 입력 이미지 전체를 대표할 수 있는 강한 특징의 feature map을 추출하게 된다.

Fc6 단계의 fully connected layer는 이전 단계의 추출된 feature map을 입력으로 받아 클래스별 분류 작업을 진행한다. 이후 활성화함수를 사용하였으며, Fc7 단계와 Fc8 단계사이 dropout 레이어 추가를 통해 일부 뉴런을 생략하여 과적합(overfitting) 문제를 보완하고, 마지막 Fc8 레이어는 카테고리 분류를 위해 softmax 함수를 사용으로 출력되는 값과의 상대적인 비교를 통해 최적의 분류 성능을 얻게된다. 이후 학습 모델 최적화를 위한 세부 파라미터 조정을 위해 기존 학습 모델의 1000개의 클래스를 인식하도록 구성된 fully connected layer의 마지막 Fc8 파라미터를 Yale Face Database B의 38명에 대한 카테고리 수에 맞추기 위해 num_output 값을 38로 수정하였다.

IV. 실험 및 결과

4.1 학습 데이터베이스 구성

얼굴인식을 위한 데이터베이스 Yale Face

Database B의 구성은 배경을 제외한 얼굴 전면부로 구성된 225×225 크기의 이미지를 총 38명에 대한 각 얼굴 영상에 대해 64가지 조명 각도에 따른 조명변화가 포함된 이미지이며, 총 2,432장으로 구성된다. 얼굴 영상은 조명 각도에 따라 5개의 Subset로 나누어지며, Subset의 예시는 그림 2과 같다.

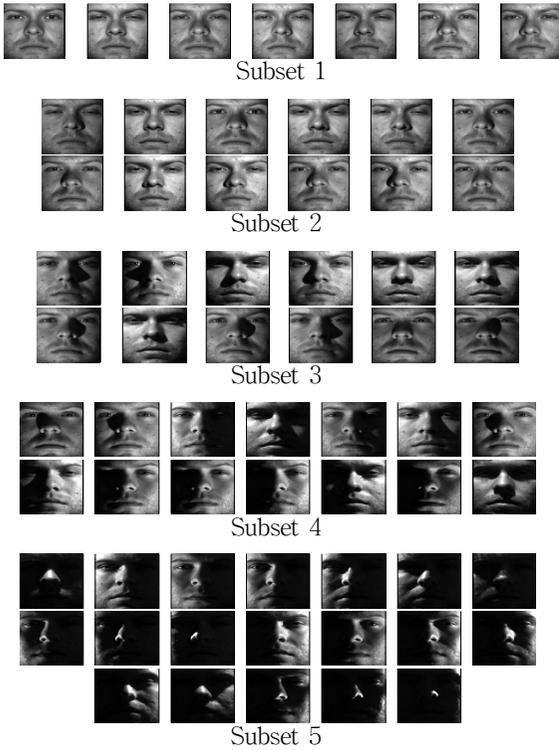


그림 2. 64가지 조명상태의 Yale Face Database B
Fig. 2 64 illumination conditions Yale Face Database B

Subset1은 (angle 0° ~ 12°)범위의 이미지 266(38명×7장)장, Subset2 (angle 13° ~ 25°) 456(38명×12장)장, Subset3 (angle 26° ~ 50°) 456(38명×12장)장, Subset4 (angle 51° ~ 77°) 532(38명×12장)장, Subset5 (angle > 77°)로 구성된다.

4.2 실험결과 및 성능평가

본 연구에서 제안하는 학습모델을 이용한 38명의 얼굴인식 성능을 확인하기 위해 각 인물 당 임의의 이름을 지정해주었으며, 1900장(38명×50장)의 학습이

미지와 342(38명×9장)장의 검증 이미지를 통해 학습을 진행하였다. 성능 테스트를 위한 데이터셋은 각 클래스 당 5장씩 추출된 190(38명×5장)장의 이미지를 통해 진행하였다. 각 Subset에 대한 인식 성능결과는 표 1과 같으며, 5장의 테스트 데이터셋을 모두 검출하는 항목의 결과는 생략하였다.

표 1. 테스트 데이터셋의 인식결과
Table 1. Recognition result of Test Dataset

Class	Sub set1 [%]	Sub set2 [%]	Sub set3 [%]	Sub set4 [%]	Sub set5 [%]
Yale 3	5	5	5	5	4
Yale 17	5	5	5	5	4
Yale 22	5	5	5	4	5
Yale 32	5	5	5	4	3
Yale 33	5	5	5	5	4

표 2. Yale Face Database B를 이용해 제안한 네트워크 성능결과

Table 2. Proposed Network of results using Yale Face Database B

Algorithm	Accuracy					
	Sub set1 [%]	Sub set2 [%]	Sub set3 [%]	Sub set4 [%]	Sub set5 [%]	Average [%]
Proposed	100	100	100	98.94	97.36	99.26

테스트결과 Subset1에 해당하는 성능결과는 100%를 기록하였다. Subset2에서는 Subset1과 같은 성능인 100%를 기록하였으며, Subset3 또한 100%를 달성했다. Subset4의 성능결과는 98.94%로 Subset3에 비해 1.06% 인식성능이 감소하였으며, Yale 3, Yale 4 2개의 이미지를 오 인식하는 결과를 보였다. Subset5의 성능결과에서는 Yale3, Yale32 2장, Yale33, Yale17 5장의 오 인식으로 Subset4보다 1.58% 감소한 97.36%의 성능결과를 보이며 조명변화 정도에 따라 얼굴인식률이 떨어지는 것을 확인할 수 있었다. 각 Subset에 대한 최종 인식률에 대한 결과는 표2와 같다. 이후 성능평가를 위해 기존의 handcraft 방식의 얼굴인식 알고리즘인 LBP, PCA, DCT와 동일한 데이터베이스로 연구된 얼굴인식 논문의 성능결과를 통해 표3과 같이 평가를 진행하였으며, 그림3과 같이 각각 9.83%, 11.61%, 4.55%의 성능향상을 보이며 여러

비선형 기법의 조합을 통해 스스로 학습[15]을 진행하는 딥러닝의 우수성을 검증하였다.

표 3. Yale Face Database B를 이용한 성능비교
Table 3. Comparison of results using Yale Face Database B

Algorithm	Accuracy					
	Sub set1 [%]	Sub set2 [%]	Sub set3 [%]	Sub set4 [%]	Sub set5 [%]	Average [%]
LBP[9]	90.11	100	83.52	85.88	78.75	87.65
PCA[8]	92.30	91.49	88.66	88.12	86.61	89.43
DCT[11]	100	100	99.17	83.57	90.83	94.71
Proposed	100	100	100	98.94	97.36	99.26

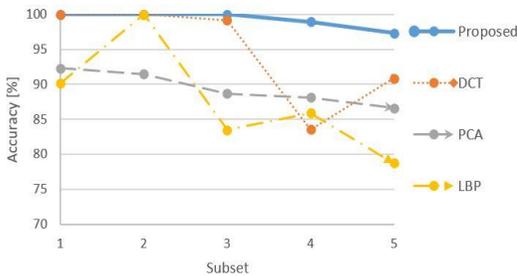


그림 3. Yale Face Database B를 이용한 알고리즘 성능 비교

Fig. 3 Graphical representation of recognition rate at Yale Face Database B using different algorithm

V. 결 론

본 연구에서는 VGG-16 모델을 기반으로한 네트워크 제작을 통해 조명 변화에 강인한 얼굴인식 모델을 제안하였다. 이 학습모델은 Yale Face Database B의 38명의 얼굴인식을 위해 최적화된 모델이며, 조명 변화에 강인한 모습을 보인다. 실험결과 190(38명×5장)장에 대한 최종 인식률은 99.26%이다. 이 수치는 기존의 얼굴인식 알고리즘인 LBP, PCA, DCT의 결과 비교를 통해 각각 평균 11.6%, 9.82%, 4.54%의 성능 향상을 보이며 CNN을 이용한 얼굴인식 방법의 적합성을 검증하였다. 추후 학습데이터의 증가를 통해 얼굴 특징의 다양성을 확보한다면 더욱 높은 인식률을 보여줄 것으로 기대된다.

감사의 글

이 논문은 2016년 순천대학교 학술연구비로 연구되었음.

References

- [1] L. Yann, B. Yoshua, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, 2015, pp. 436-444.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 21, no. 1, 2012, pp. 1097-1105.
- [3] D. Gowe, "Distinctive image features from scale-invariant key points," *Int. J. of Computer Vision*, vol. 60, no. 2, 2004, pp. 91-110.
- [4] H. Bay, T. Tuytelaars, and L. Vanl, "Surf: Speeded up robust features," *European Conf. on Computer Vision*, Heidelberg, Germany, May 2006, pp. 404-417.
- [5] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: An open source convolutional architecture for fast feature embedding," In *Proc. the 22nd ACM Int. Conf. on Multimedia*, Orlando, USA, Nov. 2014, pp. 675-678.
- [6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Proc. of Int. Conf. on Learning Representations*, San Diego, USA, Jun, 2015 pp. 1-14,
- [7] A. Georghiadis, P. Nelhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, 2001, pp. 643-660.
- [8] L. Luo, M. Swamy, and E. Plotkin, "A modified PCA algorithm for face recognition," In *Canadian Conf. on Electrical and Computer Engineering*, Quebec, Canada, vol. 1, May 2003, pp. 57-60.
- [9] J. Lee, "A Study on Face Recognition System Using LDA and SVM," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 10, no. 11,

2015, pp. 1307-1314.

- [10] A. Petpon and S. Srisuk, "Face Recognition with Local Line Binary Pattern," *In Int. Conf. on Image and Graphics*, Shanxi, China, vol. 9, Mar. 2009, pp. 533-539.
- [11] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: application to face recognition," *IEEE Trans Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, 2006, pp. 2037-2041.
- [12] X. Xiaohua, Z. Wei-shi, L. Jianhuang, and Y. Pong C, "Face illumination normalization on large and small scale features," *In IEEE Conf. on Computer Vision and Pattern Recognition*, Anchorage, USA, 2008, pp. 1-5.
- [13] H. Lien and A. Jain. "Face modeling for recognition," *In Proc. of the Int. Conf. on Image Processing*, Thessaloniki, Greece, vol. 2, Oct. 2001, pp. 693-696.
- [14] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *In Proc the IEEE*, vol. 86, no.11, 1998, pp. 2278-2324.
- [15] Y. Lee, "A Comparison and Analysis of Deep Learning Framework," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 12, no. 1, 2017, pp. 115-122.



김도연(Do-Yeon Kim)

1986년 충남대학교 계산통계학과 졸업(이학사)

2000년 충남대학교 대학원 정보통신공학과 졸업(공학석사)

2003년 충남대학교 대학원 컴퓨터공학과 졸업(공학박사)

1986 ~ 1996년 한국원자력연구원 선임연구원

1997 ~ 2008년 한국전력기술(주) 책임연구원

2008 ~ 현재 순천대학교 컴퓨터공학과 교수

※ 관심분야 : 컴퓨터비전, 컴퓨터보안

저자 소개



김연호(Yeon-Ho Kim)

2016년 순천대학교 컴퓨터공학과 졸업(공학사)

2016년 ~ 현재 순천대학교 컴퓨터공학과 석사과정

※ 관심분야 : 영상처리, 컴퓨터비전



박성욱(Sung-Wook Park)

2012년 ~ 순천대학교 컴퓨터공학과 학사과정

※ 관심분야 : 영상처리, 컴퓨터비전