

기계 학습 알고리즘을 이용한 효과적인 대상 영역 분할

장석우¹, 이경주², 정명희*

¹안양대학교 소프트웨어학과

²승실대학교 컴퓨터학과

Effective Detection of Target Region Using a Machine Learning Algorithm

Seok-Woo Jang¹, Gyungju Lee², Myunghee Jung*

¹Department of Software, Anyang University

²Department of Computer, Soongsil University

요약 다양한 종류의 컬러 영상 콘텐츠에 포함되어 있는 사람의 얼굴 영역은 다른 사람들과 특징인을 구별해 줄 수 있는 개인의 정보에 해당하므로, 입력된 컬러 영상으로부터 가려지지 않은 사람의 얼굴 영역들을 정확하게 검출하는 작업은 매우 중요하다. 본 논문에서는 입력되는 컬러 영상으로부터 기계 학습 알고리즘 중의 하나인 딥러닝 알고리즘을 이용하여 사람의 얼굴 영역을 정확하게 검출하는 방법을 제안한다. 본 논문에서 제안된 방법에서는 먼저 RGB 색상 모델로 입력되는 영상을 YCbCr 색상 모델로 변경한 다음, 기 학습된 타원형의 피부 색상 분포 모델을 활용하여 다른 영역들은 제거하고 사람의 피부 영역만을 먼저 분할한다. 그런 다음, CNN 모델 기반의 딥러닝 알고리즘을 적용하여 이전 단계에서 검출된 피부 영역 내에서 사람의 얼굴 영역을 강인하게 검출한다. 실험 결과에서는 제안된 방법이 입력되는 다양한 컬러 영상으로부터 사람의 얼굴 영역들을 기존의 방법에 비해 보다 효율적으로 분할한다는 것을 보여준다. 본 논문에서 제안된 얼굴 영역 검출 방법은 영상 보안, 물체 인식 및 추적, 얼굴 인식 등과 같은 멀티미디어 및 형태 인식과 관련된 실제적인 응용 분야에서 매우 유용하게 활용될 것으로 기대된다.

Abstract Since the face in image content corresponds to individual information that can distinguish a specific person from other people, it is important to accurately detect faces not hidden in an image. In this paper, we propose a method to accurately detect a face from input images using a deep learning algorithm, which is one of the machine learning methods. In the proposed method, image input via the red-green-blue (RGB) color model is first changed to the luminance-chroma: blue-chroma: red-chroma (YCbCr) color model; then, other regions are removed using the learned skin color model, and only the skin regions are segmented. A CNN model-based deep learning algorithm is then applied to robustly detect only the face region from the input image. Experimental results show that the proposed method more efficiently segments facial regions from input images. The proposed face area-detection method is expected to be useful in practical applications related to multimedia and shape recognition.

Keywords : Preprocessing, Color space, Machine learning, Skin area, Object region

1. 서론

최근 들어, 초고속 유무선 인터넷의 확산 및 보급, 소형화되고 저렴한 대용량 저장 장치의 발달, 고성능의 개

인화된 스마트 정보 기기의 보급으로 인해서 사람들은 인터넷을 통해 다양한 종류의 동영상 콘텐츠를 획득하기가 매우 용이해졌다[1,13]. 그리고 이런 동영상 콘텐츠는 스마트 홈(smart home), 혼합 현실(mixed reality), 게임,

*Corresponding Author : Myunghee Jung(Anyang Univ.)

Tel: +82-31-467-0963 email: mhjung@anyang.ac.kr

Received April 26, 2018

Revised May 3, 2018

Accepted May 4, 2018

Published May 31, 2018

영상 보안(security), 인공지능, 물체 인식 및 추적 등과 같은 관련된 많은 분야에서 매우 유용하게 활용되고 있다[2,12,14-16,19].

그러나 실생활에서 유용하고 편리하게 사용될 수 있는 다양한 영상 콘텐츠 서비스가 발달함에 따라서 다른 사람들과 나를 구별해 주는데 사용될 수 있는 개인 정보가 노출된 영상 빅 데이터도 기하급수적으로 증가하고 있는 실정이다. 예를 들어, 인터넷 포털의 신개념 매핑 지도 서비스 중의 하나인 거리뷰(street view) 영상[3]이 일반인에게 유용하게 사용되고 있다. 그러나 이런 거리뷰 지도 서비스가 특정한 사람의 확대된 얼굴이나 주소, 또는 전화번호 등을 가리지 않고 그대로 보여주는 등 지나치게 사적인 모습까지 담아내는데 대해 일각에서는 사생활 침해라는 우려를 표시하고 있다.

따라서 다양한 종류의 정지 또는 동영상 데이터를 분석하여 입력된 영상 내에 포함된 사람의 얼굴 영역을 강건하게 검출하는 작업은 매우 중요하다[3]. 다시 말해, 영상에 포함된 사람의 얼굴 영역은 입력 영상에 특정한 의미를 부여할 수 있는 중요한 개인 정보를 제공할 수 있으므로, 영상으로부터 개인 정보에 해당하는 사람의 얼굴 영역을 정확히 찾아내고, 이를 효과적으로 블로킹하는 연구가 필요하다.

다양한 종류의 입력 영상으로부터 사람의 얼굴 영역만을 자동으로 추출하는 기존의 연구는 관련된 참고문헌에서 찾아볼 수 있다. [4]에서는 사전에 타원형의 피부 색상 분포 모델을 정의한 다음, 이 모델을 이용하여 입력된 컬러 영상으로부터 피부 색상 영역을 추출하였다. 그런 다음, 추출된 피부 영역 내에서 눈 맵(eye map)을 이용하여 눈 영역을 검출하고, 입술 맵(lip map)을 이용하여 입 영역을 검출함으로써 사람의 얼굴 영역을 추출하였다. [5]에서는 얼굴의 색상 정보를 이용하여 대상 영상에서 부 윈도우를 효율적으로 검색하고, 애이다부스트(adaboost) 얼굴 검출기의 첫 번째 단계에 계산 속도가 매우 빠른 얼굴 색상을 이용한 얼굴 분류기를 채용함으로써 보다 빠른 얼굴 검출 속도와 보다 낮은 오검출율을 달성할 수 있는 단계별 얼굴 검출 모델을 제안하였다. [6]에서는 온라인 학습(online learning)과 기하학적인 위치 관계를 이용하여 얼굴의 구성요소인 눈, 코, 입을 검출하였으며, 이를 기반으로 사람의 얼굴 영역을 추출하였다. 이 방법은 분리된 학습 단계가 필요 없이 제안된 알고리즘이 진행되는 동안 학습이 점점 더 많이 진행된

다는 장점을 가지고 있다. [7]에서는 얼굴과 비 얼굴의 분류, 얼굴의 자세 측정, 그리고 얼굴 랜드마크 지역화를 위해서 3개의 출력 브랜치를 가진 깊은 CNN이 제안되었다. 이 모델은 각각 최대 풀링(pooling) 레이어가 뒤이어 따르는 3 개의 컨볼루션(convolution) 레이어가 있으며, 마지막 풀링 레이어 다음에는 앞에서 언급한 세 가지 브랜치의 입력을 포함하는 완전히 연결된 레이어가 있다. 위에서 기술한 방법들 이외에도 사람의 얼굴 영역 검출에 관련된 많은 서로 다른 방법들이 계속해서 관련 문헌들에 소개되고 있다[8].

위에서 기술한 기존의 얼굴 영역 검출 방법들도 어느 정도는 정확하게 사람의 얼굴 영역들을 검출할 수는 있다. 그러나 기존의 방법들은 사전에 정해진 일정한 환경 내에서 그 정확성을 보장할 수 있으며, 알고리즘의 효율적인 측면에서는 다소 개선이 필요하다는 단점이 존재한다[17, 18].

따라서 본 논문에서는 기계 학습 알고리즘을 이용하여 입력 영상으로부터 얼굴 영역만을 정확하고 효율적으로 검출해 내는 새로운 알고리즘을 제안한다. 아래 Fig. 1은 본 논문에서 제안하는 얼굴 영역 검출 방법의 전체적인 개요도를 보여준다.

Fig. 1에서 보는 바와 같이 본 논문에서 제안된 알고리즘에서는 먼저 입력된 영상의 RGB 색상 모델을 $YCbCr$ 색상 모델로 변경한 다음, 사전에 학습된 타원형의 피부 색상 분포 모델을 사용하여 배경과 같은 다른 영역들은 제거하고 사람의 피부 색상 분포 영역만을 입력된 영상으로부터 분할한다. 그런 다음, CNN(convolutional neural network) 모델[9] 기반의 딥러닝 알고리즘을 적용하여 검출된 피부 영역으로부터 사람의 얼굴 영역만을 강인하게 검출한다.

1장에서는 본 연구를 수행하게 된 개요에 대해 설명하였다. 2장에서는 사람의 피부 색상 영역을 학습하고, 이를 강인하게 검출하는 방법에 대해 기술한다. 3장에서는 기계 학습 알고리즘을 이용하여 입력된 컬러 영상으로부터 사람의 얼굴 영역을 강인하고 효율적으로 검출하는 방법에 대해 설명한다. 그리고 4장에서는 제안된 얼굴 영역 검출 알고리즘의 성능을 비교 평가하기 위한 실험결과를 보이며, 5장에서는 결론 및 향후 연구방향을 기술한다.

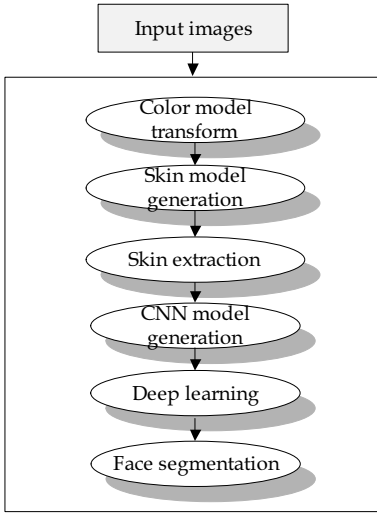


Fig. 1. Overall flow of the proposed algorithm

2. 피부 영역 검출

본 장에서는 입력된 컬러 영상들을 분석하여 배경과 같은 다른 영역들은 제외하고 사람의 피부 색상 분포 영역들만을 검출하는 방법에 대해 기술한다. 이를 위해, 먼저 식 (1)을 이용하여 RGB 컬러 공간으로 입력되는 영상을 YC_bC_r 컬러 공간의 영상으로 변환한다. 일반적으로 YC_bC_r 컬러 공간이 다른 컬러 공간들에 비해서 사람의 피부 영역을 검출하기가 보다 효과적이라고 알려져 있다. 그런 다음, 타원형 모델(elliptical model)을 기반으로 생성한 피부 색상 분포 모델을 사용해 입력된 컬러 영상으로부터 사람의 피부 색상 영역들을 강인하게 검출한다[4].

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.533 & 24.966 \\ -37.797 & -74.203 & 112 \\ 122 & -93.786 & -84.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

본 논문에서는 YC_bC_r 색상 공간의 C_bC_r 공간에서 타원형의 피부 색상 분포 모델을 정의하지 않고, 색상(chroma) 정보가 명도(luma) 정보에 최대한 독립적이 되도록 구성하기 위해서 아래의 식 (2), (3), (4), (5)를 활용하여 변환된 $C_b'C_r'$ 공간에서 피부 색상 모델을 정의한다. 보통 YC_bC_r 컬러 공간에서 색상 정보인 C_b 와 C_r 은 명도 정보인 Y 의 함수 $C_b(Y)$ 와 $C_r(Y)$ 로 표현될 수 있다.

변환된 색상 정보를 $C_b'(Y)$ 와 $C_r'(Y)$ 로 표현해 보자. 그리고 타원형 피부 색상 모델은 중심($\bar{C}_b(Y)$, $\bar{C}_r(Y)$)과 클러스터의 퍼짐($W_{cb}(Y)$, $W_{cr}(Y)$)으로 지정될 수 있으며, 이 값들은 아래의 식 (2)와 같이 변환되는 색상인 C_b' 과 C_r' 을 계산하기 위해 사용된다.

$$C'_i(Y) = \begin{cases} (C_i(Y) - \bar{C}_i(Y)) \frac{W_{c_i}}{W_{C_i}(Y)} + \bar{C}_i(Y) & \text{if } Y < K_l \text{ or } K_h < Y \\ C_i(Y) & \text{if } Y \in [K_l, K_h] \end{cases} \quad (2)$$

$$W_{C_i}(Y) = \begin{cases} WL_{C_i} + \frac{(Y - Y_{\min})(W_{C_i} - WL_{C_i})}{K_l - Y_{\min}} & \text{if } Y < K_l \\ WH_{C_i} + \frac{(Y_{\max} - Y)(W_{C_i} - WH_{C_i})}{Y_{\max} - K_h} & \text{if } K_h < Y \end{cases} \quad (3)$$

$$\bar{C}_b(Y) = \begin{cases} 108 + \frac{(K_l - Y)(118 - 108)}{K_l - Y_{\min}} & \text{if } Y < K_l \\ 108 + \frac{(Y - K_h)(118 - 108)}{Y_{\max} - K_h} & \text{if } K_h < Y \end{cases} \quad (4)$$

$$\bar{C}_r(Y) = \begin{cases} 154 - \frac{(K_l - Y)(154 - 144)}{K_l - Y_{\min}} & \text{if } Y < K_l \\ 154 + \frac{(Y - K_h)(154 - 132)}{Y_{\max} - K_h} & \text{if } K_h < Y \end{cases} \quad (5)$$

여기에서 C_i 는 C_b 또는 C_r 를 나타내며, $W_{cb}=46.97$, $WL_{cb}=23$, $WH_{cb}=14$, $W_{cr}=38.76$, $WL_{cr}=20$, $WH_{cr}=10$, $K_l=125$, 그리고 $K_h=188$ 을 의미한다. 이 파라미터들은 HHI(Heinrich-Hertz-Institute) 영상 데이터 집합의 피부 패치(patch) 샘플을 활용하여 학습 단계에서 추출되었다. 그리고 YC_bC_r 컬러 공간의 Y_{\min} 과 Y_{\max} 는 각각 16과 235이다. 그리고 변환된 $C_b'C_r'$ 공간에서 피부 톤(tone)에 대한 타원형 모델은 식 (6)과 식 (7)로 표현될 수 있다.

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} = 1 \quad (6)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} C'_b - c_x \\ C'_r - c_y \end{bmatrix} \quad (7)$$

식 (6)과 (7)에서 $c_x=109.38$, $c_y=152.02$, $\Theta=2.53$ (라디안), $ec_x=1.60$, $ec_y=2.41$, $a=25.39$, $b=14.03$ 이며, 이 값들은 $C_b'C'_a$ 공간에서 군집화된 피부 색상 영역으로부터 산출된 값이다. 이와 같이 변환된 $C_b'C'_a$ 공간에서 정의된 피부 색상 모델은 변환되지 않은 C_bC_r 공간에서 정의된 피부 색상 모델에 비해 보다 넓은 범위의 명도 값을 가지는 피부 색상 화소에 대한 검출이 가능하다.

다음으로는 검출된 피부 색상 분포 영역 내에 존재하는 잡음(noise)들을 제거하기 위해서 모폴로지 연산(morphological operation)을 수행한 다음, 레이블링을 수행하여 최종적인 피부 영역을 검출한다. 그러면 기계 학습을 이용한 얼굴 영역 검출을 위한 관심 영역 추출이 완료된다.

3. 기계 학습을 이용한 얼굴 검출

본 논문에서는 이전 단계에서 검출한 피부 색상 화소들을 대상으로 기계 학습 알고리즘[10] 중의 하나인 CNN(convolutional neural network)을 이용하여 사람의 얼굴 영역들을 검출한다. 보통 CNN 모델은 영상의 2차원 데이터를 입력으로 활용할 수 있으므로 영상 내의 위치 관련 특징들이 유실되지 않는다. CNN은 컨볼루션(convolution) 레이어와 풀링 레이어(pooling layer)를 통해 2차원 데이터로부터 특징을 추출하며, 이 특징들은 보통의 다층 인공 신경망에 입력된다. Fig. 2는 CNN의 기본적인 구조를 그림으로 보여준다.

CNN은 입력받은 벡터를 일련의 히든 레이어(hidden layer)를 통해 변형시킨다. 여기에서 히든 레이어는 뉴런들로 이루어져 있으며, 각 뉴런은 앞쪽 레이어(previous layer)의 모든 뉴런과 연결되어 있다. 그리고 같은 레이어 내에 있는 뉴런들끼리는 연결이 존재하지 않고 서로 독립적이다. 마지막으로, 전체 연결(fully-connected) 레이어는 출력 레이어라고 불리우며, 분류 문제에서 클래스

스 점수(class score)를 나타낸다.

보통 일반적인 신경망은 영상을 다루기가 적절하지 않다고 알려져 있다. 예를 들어, 영상이 $32 \times 32 \times 3$ ($W \times H \times C$)으로 이루어진 경우라면 첫 번째 히든 레이어 내의 하나의 뉴런에는 $32 \times 32 \times 3=3,072$ 개의 가중치가 필요하다. 만일 보다 더 큰 이미지를 사용할 경우에는 이러한 신경망 구조를 이용하는 것은 어렵다. 다시 말해, $200 \times 200 \times 3$ 의 영상에는 같은 뉴런에 대해 120,000개의 가중치가 필요하기 때문이다. 더욱이 이런 뉴런들이 레이어 내에 여러 개가 존재하므로 모수의 개수가 크게 증가하게 된다. 이와 같이 전체 연결 형태는 심한 낭비이며, 많은 수의 모수는 결국 오버피팅(overfitting)으로 귀결하게 된다.

CNN은 입력이 영상으로 이루어져 있다는 특징을 살려 좀 더 합리적인 방향으로 신경망의 구조를 구성할 수 있다. 다시 말해, 일반 신경망과는 달리 CNN의 레이어들은 가로, 세로, 깊이(채널)의 3개 차원을 가진다. 하나의 레이어에 위치한 뉴런들은 일반적인 신경망과는 달리 앞 레이어의 전체 뉴런들이 아닌 일부에만 연결이 되어 있다. 그리고 CNN의 구조는 전체 영상을 클래스 점수들로 이루어진 하나의 벡터로 만들어 주기 때문에 마지막 출력 레이어는 $1 \times 1 \times N$ 의 차원을 가지게 된다. 여기에서 N은 클래스의 개수를 나타낸다.

CNN에서 컨볼루션 레이어는 합성곱 연산을 통해 특징들을 추출하는 레이어로, 규칙적인 패턴을 가진 곱셈이 이루어진다. 즉, 컨볼루션 레이어는 커널(kernel)이라고 불리는 필터를 적용하는 레이어이다. 여기에서, 합성곱한 영상은 특징맵이라고 불리운다. 예를 들어, 20×20 입력 영상에서 5×5 의 패치를 이용해 컨볼루션 레이어를 구축하면 이 레이어의 특징맵은 16×16 으로 생성된다. CNN의 장점 중의 하나는 학습 알고리즘을 이용해 스스로 적당한 커널 값을 학습함으로써 합성곱 레이어의 필터 값들을 손으로 직접 설정할 필요가

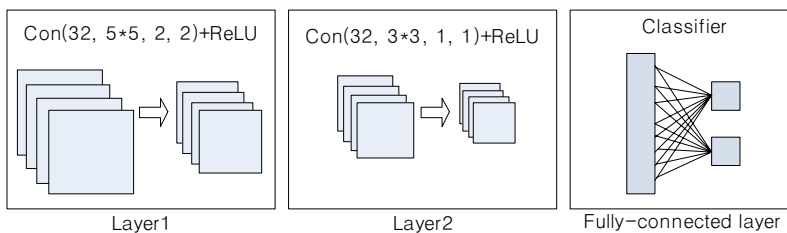


Fig. 2. Structure of convolutional neural network

없다는 사실이다.

보통 신경망에서 활성화 함수(activation function)로 많이 사용되는 시그모이드(sigmoid) 함수는 식 (1)과 같이 정의되어 결과가 0에서 1사이의 값으로 출력된다. 시그모이드 함수는 뇌의 뉴런과 유사하여 많이 사용되었으나, 신경망의 층이 깊어질수록 그래디언트(gradient)가 0으로 수렴하는 문제가 발생하여 딥러닝의 경우에는 잘 사용되지 않는다. 예를 들어, 히든 레이어의 개수가 9개에서 10개 정도가 되면 학습이 잘 진행되지 않는다.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (8)$$

반면에 ReLU는 최근 신경망에서 많이 사용되는 활성화 함수로, 계산이 매우 효율적이고 수렴 속도가 시그모이드 함수에 비해 약 6배 정도 빠르다는 장점이 있어 본 논문에서도 ReLU를 사용한다.

$$f(x) = \max(0, x) \quad (9)$$

CNN에서 풀링 레이어는 학습을 수행하지 않으며, 컨볼루션 레이어에서 전파된 특징의 차원을 축소(다운샘플링)하여 모델의 복잡도를 줄이는 역할을 수행함으로써 입력 데이터로부터 강건한 특징을 추출할 수 있다. Fig. 3은 최대 풀링(max pooling)과 평균 풀링(average pooling)의 예시를 보여준다. Fig. 3에서 컨볼루션 레이어를 통해 4×4의 특징맵이 생성되었다면, 풀링을 통해 2×2로 차원이 축소된다. 이때 최대 풀링은 4×4 영역에서 2×2씩 영역을 분할하여 각 영역의 최대값을 이용하는 방식이고, 평균 풀링은 평균값으로 대체하는 방식이다.

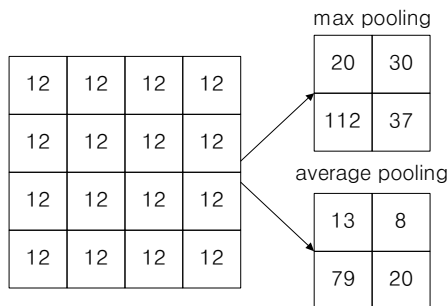


Fig. 3. Example of Pooling

CNN에서 전체 연결 레이어는 앞의 과정에서 추출된 특징을 신경망에 넣어서 분류하는 과정을 거치는 단계이며, N개의 클래스로 분류한다.

4. 실험 결과

본 논문에서 실험을 위해서 사용한 컴퓨터는 인텔 Core(TM) i7-6700 3.4Ghz의 CPU와 16GB의 메인 메모리, 256GB인 SSD, 그리고 NVIDIA의 GPU GP104가 사용된 갤럭시 Geforce GTX 1080 Ti 그래픽 카드로 구성되어 있으며, 윈도우 10 운영체제(OS)를 이용하였다. 그리고 통합 개발 환경으로는 마이크로소프트사의 비주얼 스튜디오(Visual Studio) 버전 2015를 사용하였으며, OpenCV 오픈 라이브러리 및 Dlib C++ 라이브러리를 이용하여 제안된 알고리즘을 구현하였다.

본 논문에서 얼굴 검출 실험을 위한 영상 데이터로는 영국 옥스퍼드 대학의 VGG에서 제작한 얼굴 데이터셋 [11]를 사용하였다. 딥러닝을 위한 학습 데이터로는 5,000장의 영상을 사용하였고, 얼굴 검출을 위한 테스트 데이터로는 500장의 영상을 사용하였다.

본 논문에서 사용한 딥러닝의 기본적인 구조는 Fig. 2와 같으며, 학습률(learning rate)은 0.1에서 시작하여 0.0001까지 순차적으로 감소시켜 사용하였다.

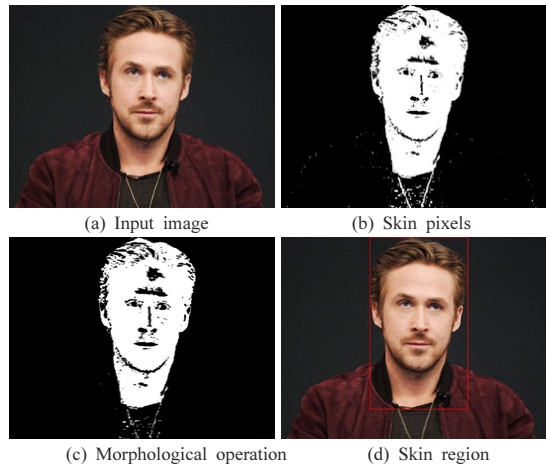


Fig. 4. Detected skin region

Fig. 4 (a)는 본 논문에서 사용한 사람의 얼굴 영역이 포함된 입력된 컬러 영상의 한 예를 보여준다. Fig. 4 (b)

는 입력된 컬러 영상으로부터 타원형의 피부 색상 분포 모델을 이용하여 피부 색상 화소들을 추출한 결과를 보여준다. Fig. 4 (c)는 추출된 피부 색상 화소들에 모폴로지 연산(morphological operation)을 수행하여 잡음들을 제거한 후의 영상을 보여준다. 그리고 Fig. 4 (d)는 입력 영상에 검출한 피부 색상 분포 영역만을 표시하는 최소 포함 사각형(MER: minimum enclosing rectangle)을 표시한 결과를 보여준다.

본 논문에서는 제안된 사람의 얼굴 영역 검출 방법의 정확도 및 속도 측면에서의 성능을 정량적으로 비교 및 평가하였다. 비 얼굴 영역을 포함하지 않은 정확하게 검출된 사람의 얼굴 영역들과 영상에 존재하는 전체 얼굴 영역들과의 비율을 백분율(%)로 표시한 식 (7)과 같은 정확도 척도를 사용하였다. 식 (7)에서 $F_{detected}$ 는 제안된 알고리즘을 이용하여 정확하게 검출한 얼굴 영역들의 개수를 나타내고, F_{total} 은 테스트 영상에 존재하는 얼굴 영역들의 전체 개수를 나타내며, 정확도 척도 $R_{accuracy}$ 는 백분율로 표시된다.

$$R_{accuracy} = \frac{F_{detected}}{F_{existing}} \times 100 (\%) \quad (7)$$

정확도 측면에서는 VGG 영상 300장을 가지고 실험을 하였다. 실험 결과 전처리를 수행하지 않은 기존의 방법이 93.55%의 성능을 보였으며, 피부 검출을 수행하는 제안된 방법이 92.12%의 성능을 보였다. 다시 말해, 제안된 방법이 다소 낮은 성능을 보였으나 전반적으로 근소한 차이를 보였다.

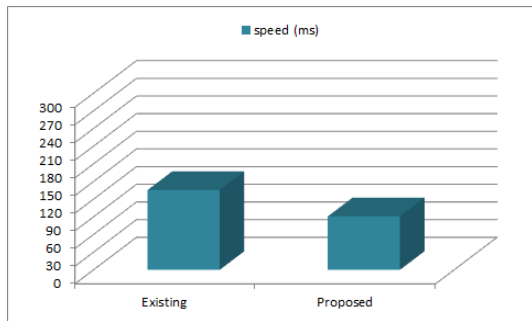


Fig. 5. Performance evaluation

그리고 속도 측면에서의 성능은 각 방법으로 얼굴 검출이 수행되는 평균적인 시간을 측정하였다. 속도 측정

을 위해 영상의 평균 해상도는 463×413을 사용하였고, 윈도우의 크기는 40×40, 그리고 윈도우는 2 화소 또는 1 화소씩 이동하였다. 또한, 6단계의 영상 피라미드를 생성하여 CNN의 입력으로 사용하였다. Fig. 5는 얼굴 영역 검출 알고리즘의 속도 측면의 성능 비교 결과를 그래프로 나타내고 있다. Fig. 5에서 확인할 수 있듯이 제안된 알고리즘을 이용한 방법이 보다 빠르게 얼굴 영역을 검출한다는 것을 확인할 수 있었다. 다시 말해, 피부 영역을 검출하지 않고 입력 영상에 대해 직접 얼굴 검출을 수행하는 기존의 방법은 입력 영상 전체에 대해 얼굴 검출을 시도하므로 시간이 오래 소요되었다. 그러나 제안된 방법은 전처리를 통해 피부 영역을 먼저 검출한 다음, 검출된 피부 영역 내에서만 얼굴 검출을 시도하였으므로 보다 빠르게 얼굴 영역을 검출할 수 있었다.

실험 결과를 전체적으로 살펴보면, 실제 얼굴 영상을 얼굴이 아니라고 오 분류한 결과를 보면, 모자, 선글라스 등으로 인한 얼굴의 폐쇄, 얼굴의 회전, 그리고 색상 있는 조명과 잡음의 영향으로 인해 사람의 피부 색상이 왜곡되어 촬영된 영상들이 많았다. 또한 얼굴이 포함되어 있지 않은 영상을 얼굴로 분류한 결과를 살펴보면 피부 색상과 비슷한 물체나 배경이 존재하는 경우가 많았다.

5. 결론

최근 들어, 노출된 개인 정보를 보호할 목적으로, 입력된 서로 다른 컬러 영상들을 분석하여 사람의 얼굴 영역을 자동으로 검출하는 연구에 대한 필요성이 점점 더 증가하고 있다.

본 논문에서는 입력되는 컬러 영상으로부터 기계 학습 방법 중의 하나인 딥러닝 알고리즘을 이용하여 사람의 얼굴 영역을 정확하게 검출하는 방법을 제안하였다. 제안된 방법에서는 먼저 RGB 색상 모델로 입력되는 영상을 $YCbCr$ 색상 모델로 변경한 다음, 사전에 학습된 피부 색상 분포 모델을 활용하여 다른 영역들은 제거하고 사람의 피부 영역만을 분할하였다. 그런 다음, CNN 모델 기반의 딥러닝 알고리즘을 적용하여 입력 영상으로부터 얼굴 영역만을 강인하게 검출하였다. 실험 결과에서는 제안된 방법이 입력되는 다양한 컬러 영상으로부터 사람의 얼굴 영역들을 보다 효율적으로 분할한다는 것을 보여주었다.

향후에는 본 논문에서 제안된 얼굴 영역 검출 방법을 보다 다양한 종류의 환경에서 촬영된 입력 영상들에 적용하여 제안된 방법의 정확도 및 속도 측면에서의 성능을 비교 평가할 계획이다. 그리고 현재까지 진행된 얼굴 영역 검출 알고리즘에 얼굴 인식 알고리즘을 추가할 예정이다.

References

- 6, pp. 822-833, 2017.
DOI: <https://doi.org/10.1109/THMS.2017.2700630>
- [11] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," *In Proc. of the 26th British Machine Vision Conference*, pp. 1-12, September 2015.
DOI: <https://doi.org/10.5244/C.29.41>
- [12] Jae-Ho Bae, Development of Dynamic Magnetic Field Emulator for Smart Multi-Card, *J. Soc. Korea Ind. Syst. Eng.*, Vol. 40, No. 4, pp. 183-190, Dec. 2017.
DOI: <https://doi.org/10.11627/jkise.2017.40.4.183>
- [13] Joong-Bae Lee, Dong-Hyun Baek, The Effect of Smartphone Purchasing Determinants on Repurchase Intention, *J. Soc. Korea Ind. Syst. Eng.*, Vol. 40, No. 2, pp. 1-12, June 2017.
DOI: <https://doi.org/10.11627/jkise.2017.40.2.001>
- [14] Hee-Ohl Kim, Dong-Hyun Baek, A Study on Categorization of Accident Pattern for Organization's Information Security Strategy Establish, *J. Soc. Korea Ind. Syst. Eng.*, Vol. 38, No. 4, pp. 193-201, Dec. 2015.
DOI: <http://dx.doi.org/10.11627/jkise.2015.38.4.193>
- [15] Hee-Ohl Kim, Dong-Hyun Baek, Prioritize Security Strategy based on Enterprise Type Classification Using Pair Comparison, *J. Soc. Korea Ind. Syst. Eng.*, Vol. 39, No. 4, pp. 97-105, Dec. 2016.
DOI: <http://dx.doi.org/10.11627/jkise.2016.39.4.097>
- [16] Ji Hoon Kyung, Chong Su Kim, A Study on Measurements of IT Security Service Quality : Feasibility of Quantitative Measures, *J. Soc. Korea Ind. Syst. Eng.*, Vol. 38, No. 4, pp. 30-38, Dec. 2015.
DOI: <http://dx.doi.org/10.11627/jkise.2015.38.4.30>
- [17] Tae-Hyun Ahn, Jae-Gyun Park, Young-Man Kwon, A Study on Performance of ML Algorithms and Feature Extraction to detect Malware, *The Journal of The Institute of Internet, Broadcasting and Communication (IIBC)*, Vol. 18, No. 1, pp.211-216, Feb. 28, 2018.
DOI: <http://doi.org/10.7236/JIIBC.2018.18.1.211>
- [18] Seung-Jae Kim, Jung-Jae Lee, A Study on Face Recognition using Support Vector Machine, *The Journal of The Institute of Internet, Broadcasting and Communication (IIBC)*, Vol. 16, No. 6, pp. 183-190, Dec. 2016.
DOI: <https://doi.org/10.7236/JIIBC.2016.16.6.183>
- [19] Min-soo Kang, Chunhwa Ihm, Jaeyeon Lee, Eun-Hye Choi, Sang Kwang Lee, A Study on Object Recognition for Safe Operation of Hospital Logistics Robot Based on IoT, *The Journal of The Institute of Internet, Broadcasting and Communication (IIBC)*, Vol. 17, No. 2, pp. 141-146, Apr. 2017.
DOI: <https://doi.org/10.7236/JIIBC.2017.17.2.141>
- [1] J.-G. Ko, Y.-S. Bae, J.-Y. Park, and K. Park, "Technologies Trends in Image Big Data Analysis," *Electronics and Telecommunications Trends*, Vol. 29, No. 4, pp. 21-29, August 2014.
- [2] S. Zhang, P. McCullagh, H. Zheng, and C. Nugent, "Situation Awareness Inferred from Posture Transition and Location: Derived from Smartphone and Smart home Sensors," *IEEE Transactions on Human-Machine Systems*, Vol. 47, Issue 6, pp. 814-821, 2017.
DOI: <https://doi.org/10.1109/THMS.2017.2693238>
- [3] T.-H. Tsai, W.-H. Cheng, C.-W. You, M.-C. Hu, A. W. Tsui, and H.-Y. Chi, "Learning and Recognition of On-Premise Signs From Weakly Labeled Street View Images," *IEEE Transactions on Image Processing*, Vol. 23, No. 3, pp. 1047-1059, 2014.
DOI: <https://doi.org/10.1109/TIP.2014.2298982>
- [4] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face Detection in Color Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp. 696-706, May 2002.
DOI: <https://doi.org/10.1109/34.1000242>
- [5] Y. N. Chae, J.-N. Chung, and H. S. Yang, "Efficient Face Detection Using Adaboost and Facial Color," *Journal of Korean Institute of Information Science and Engineers*, Vol. 36, No. 7, pp. 548-558, 2009.
- [6] K.-M. Lee, "Component-based Face Detection and Verification," *Pattern Recognition Letters*, Vol. 29, pp. 200-214, 2008.
DOI: <https://doi.org/10.1016/j.patrec.2007.09.013>
- [7] C. Zhang and Z. Zhang, "Improving Multiview Face Detection with Multi-Task Deep Convolutional Neural Networks," *In Proc. of the IEEE Winter Conference on Applications of Computer Vision*, pp. 1036 - 1041, 2014.
DOI: <https://doi.org/10.1109/WACV.2014.6835990>
- [8] D. Triantafyllidou, P. Nousi, and A. Tefas, "Fast Deep Convolutional Face Detection in the Wild Exploiting Hard Sample Mining," *Big Data Research*, July 2017.
- [9] L. Ren, J. Lu, J. Feng, and J. Zhou, "Multi-Modal Uniform Deep Learning for RGB-D Person Re-Identification," *Pattern Recognition*, Vol. 72, pp. 446-457, December 2017.
DOI: <https://doi.org/10.1016/j.patcog.2017.06.037>
- [10] M. Li, J. Wei, X. Zheng, and M. L. Bolton, "A Formal Machine Learning Approach to Generating Human-Machine Interfaces from Task Models," *IEEE Transactions on Human Machine Systems*, Vol. 47, No.

장 석 우(Seok-Woo Jang)

[정회원]



- 1995년 2월 : 송실대학교 전자계학과 (공학사)
- 1997년 2월 : 송실대학교 컴퓨터학과 (공학석사)
- 2000년 8월 : 송실대학교 컴퓨터학과 (공학박사)
- 2009년 3월 ~ 현재 : 안양대학교 소프트웨어학과 교수

<관심분야>

로봇비전, 증강현실, HCI, 비디오 색인 및 검색, 등

이 경 주(Gyungju Lee)

[정회원]



- 2010년 2월 : 군산대학교 컴퓨터정보공학과 (공학사)
- 2013년 2월 : 송실대학교 컴퓨터학과 (공학석사)
- 2013년 3월 : 송실대학교 컴퓨터학과 박사과정

<관심분야>

컴퓨터비전, 영상처리, 패턴인식 등

정 명 희(Myunghee Jung)

[정회원]



- 1989년 2월 : 서울대학교 계산통계학과 (공학사)
- 1991년 12월 : University of Texas, Austin, (공학석사)
- 1997년 8월 : University of Texas, Austin, (공학박사)
- 2011년 6월 ~ 현재 : 안양대학교 소프트웨어학과 교수

<관심분야>

통신, 시뮬레이션, 원격탐사 데이터 처리, 영상처리 멀티미디어 등