

심층 합성곱 신경망을 이용한 교통신호등 인식

김민기

Traffic Light Recognition Using a Deep Convolutional Neural Network

Min-Ki Kim[†]

ABSTRACT

The color of traffic light is sensitive to various illumination conditions. Especially it loses the hue information when oversaturation happens on the lighting area. This paper proposes a traffic light recognition method robust to these illumination variations. The method consists of two steps of traffic light detection and recognition. It just uses the intensity and saturation in the first step of traffic light detection. It delays the use of hue information until it reaches to the second step of recognizing the signal of traffic light. We utilized a deep learning technique in the second step. We designed a deep convolutional neural network(DCNN) which is composed of three convolutional networks and two fully connected networks. 12 video clips were used to evaluate the performance of the proposed method. Experimental results show the performance of traffic light detection reporting the precision of 93.9%, the recall of 91.6%, and the recognition accuracy of 89.4%. Considering that the maximum distance between the camera and traffic lights is 70m, the results shows that the proposed method is effective.

Key words: Traffic Light Recognition, Deep Convolutional Neural Network

1. 서론

구글(Google)의 딥마인드가 개발한 알파고(Alpha Go)라는 인공지능 바둑 프로그램의 출현을 계기로, 인공지능 기술을 이용한 다양한 장치 및 제품의 지능화가 폭넓게 진행되고 있다. 특히 자동차를 지능화하여 자율적으로 운행할 수 있도록 하는 자율주행 자동차에 대한 연구가 기존의 자동차 업계는 물론 구글을 비롯한 정보기술 업계에서 활발히 이루어지고 있다. 자율주행 자동차 개발을 위해서는 도로의 상황 인지 기술이 필수적이다. 운전자가 자동차 운행 중에 지속적으로 탐지해야 할 주요 정보들, 예를 들어 차선, 신호등, 주변의 주행 차량, 보행자 등의 정보들을 자동으로 인지할 수 있어야 한다. 신호등은 교차로의

진입 방향에 따라 순차적으로 자동차의 통행을 제어하는 신호로, 차선과는 달리 동적으로 변하는 신호이기 때문에 교차로에 진입할 때 운전자들의 주의 집중이 요구된다. 2017년 교통사고 통계분석 자료에 따르면 교차로에서 발생한 교통사고가 전체의 47.3%였고, 사망자는 33.1% 달하였다. 따라서 신호등을 자동으로 인지하여 운전자에게 정보를 제공한다면 교차로에서의 교통사고를 줄이는 데에도 기여할 수 있을 것이다.

신호등 인식은 신호등의 검출과 판별의 두 부분으로 나누어 볼 수 있다. 신호등은 대개 검은색 함체에 점멸등이 배치되어 있고, 함체는 사각형의 형태를 가지므로 기존의 연구들은 이러한 특징을 추출하여 신호등을 검출하였다. Cai 외[1]는 RGB 컬러모델을 사

※ 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 : Aug. 7, 2018, Revision date : Sep. 27, 2018
Approval date : Oct. 18, 2018

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

용하여 어둡고 채도가 낮은 영역들을 추출한 후 각 영역에 대한 면적, 중회비 등을 통해 신호등 합체를 추출하였다. 이러한 방식은 국내 연구에서도 다수 진행된 바 있다. [2-4]의 연구에서는 신호등 합체 영역에 나타나는 적색, 녹색, 황색, 흑색의 범위를 HSV 컬러모델로 정의하여 신호등 합체 및 점등 영역을 분할하였다. 본 연구에서는 신호등 합체를 탐색하지 않고 신호등 점등 영역만을 검출한 후 딥러닝 기술을 이용하여 신호등을 인식하는 방법을 제안한다. 적색, 황색, 녹색 신호등 영상이 내포하는 특징을 심층 신경망으로 학습시킨 후 도로 영상에서 관심 영역을 추출하여 신경망의 입력으로 제공하면 해당 영역이 신호등 인지, 신호등 이라면 어떤 신호를 나타내는지 식별하는 방식이다.

본 논문의 구성은 다음과 같다. 2장에서는 신호등 인식에 관련된 기존 연구들의 접근 방식을 살펴보고, 기존 방식의 장단점을 기술한다. 3장과 4장에서는 본 논문에서 제안하는 신호등 점등 영역 추출 방법과 심층 합성곱 신경망을 이용한 신호등 인식 방법을 제시한다. 그리고 블랙박스 동영상을 이용한 실험 결과를 5장에서 제시하고, 이를 통해 제안된 신호등 인식 방법의 타당성을 제시한다. 마지막으로 6장에서 결론과 향후 과제를 제시한다.

2. 관련 연구

도로 영상에서 신호등을 인식하는 접근 방식은 주야간 시간대에 따라 큰 차이가 나타난다. 야간 도로 영상은 전반적으로 어둡고 채도가 낮게 나타나며, 신호등 합체를 식별할 수 없는 경우가 다수 나타나기 때문이다. 본 연구는 주간 영상에서 신호등을 인식하는 문제를 다루고 있으므로, 주간 영상에서 신호등을 검출하고 인식하는 연구들에 대하여 살펴본다. 색상 정보는 신호등 검출은 물론 신호등 판별을 위한 필수적인 정보이다. 컬러모델에 따라 시점 및 조도 변화에 따른 색상 변화의 민감도가 다르게 나타난다[5]. 이러한 이유로 입력 컬러영상으로부터 신호등이 나타내는 색상 정보를 안정적으로 검출하기 위한 다양한 컬러모델에 대한 탐색 및 적용이 이루어져 왔다.

RGB 컬러모델은 날씨 및 시간대에 따른 조도변화에 따라 RGB 성분 값이 민감하게 변하기 때문에 Omachi 외[6]는 정규화된 RGB 컬러공간인 rgb 컬러공간에서 적색, 황색, 녹색, 검정색의 범위를 정의하

고 있으며 여기에 속하지 않는 영역은 백색으로 취급하였다. 소벨(Sobel) 연산자로 에지를 검출한 후 허프 변환(Hough transform)을 이용한 원형 검출을 통해 신호등을 검출하였다. Diaz-Cabrera 외[7]의 연구에서도 정규화된 RGB 컬러모델을 이용하고 있는데, rgb 채널별 임계값으로 색 구분을 위한 범위를 설정하지 않고 rgb 채널들 간의 차이 값을 색 구분을 위한 임계값으로 설정하였다. 적색, 녹색, 황색 영역들이 검출되면 해당 영역의 외접 사각형이 정사각형에 가까운지를 판별한다. 신호등에서 색상별 점등 영역의 위치가 고정되어 있으므로 이러한 도메인 지식을 이용하여 이웃한 영역에 나타나는 소등된 영역을 탐지하여 신호등을 인식한다. 신호등 인식에 대한 국내 연구들은 다수가 HSV 컬러모델을 이용하고 있다. 이들 연구들은 색상, 채도, 명도 채널별로 임계값을 설정한 후 영상을 구성하는 임의 화소를 설정된 범위와 비교하여 특정 영역이 신호등의 일부인지 그리고 어떤 색상인지를 판단한다. Kim 외[3]의 연구에서는 신호등 합체 영역에 나타나는 적색, 녹색, 황색, 흑색의 범위를 정의하고 있으며, Jeong 외[4]의 연구에서는 녹색과 검정색의 경우에는 (채도-명도) 값의 범위를 추가적으로 설정하였다. Jo 외[2]는 황색을 제외한 적색, 녹색, 검정색의 3가지 색상에 대한 범위를 설정하여 신호등 점등영역을 검출하였다. 그러나 이들 연구에서 설정한 색상별 임계값의 범위가 상당한 차이를 보이고 있기 때문에 제시된 색상별 임계값을 이용하는데 어려움이 있다. 색상 정보를 이용하는 연구들이 내포하는 다른 문제점은 신호등 점등 영역이 과포화(over-saturated)되어 색 정보가 유실되는 상황에 대처하기 어렵다는 것이다.

Moizumi 외[8]는 $L^*a^*b^*$ 컬러모델을 이용하여 색상과 명도를 분리하였다. 전형적인 신호등 영역을 검출할 때는 a^* 채널과 b^* 채널을 이용하고, 과포화된 신호등 영역을 검출하기 위하여 L^* 채널을 사용함으로써 적색이나 녹색 점등 영역이 과포화에 의해 색의 변화가 발생하는 문제에 대처할 수 있는 장점이 있다. Kim[9]은 채도와 명도 정보를 기반으로 점등 영역을 검출하며, 조도 변화에 강인한 정규화된 RGB 컬러모델을 이용하여 점등 영역 색상을 모델링 하였다. 적색등, 황색등, 녹색등의 3가지 색상과 소등되었을 때의 검정색, 그리고 과포화에 의한 점등 영역을 백색으로 모델링하여 신호등 인식을 수행하였다.

최근에는 SVM(support vector machine), DNN (deep neural network) 등 기계학습을 이용하여 신호등을 인식하는 연구들이 발표되고 있다. Ozcelik 외 [10]는 HSV 컬러공간에서 허프 변환을 통해 추출한 신호등 후보 영역을 SVM 분류기를 이용하여 신호등을 식별하였다. Jang 외[11]와 Zhou 외[12]는 각각 HSV와 RGB 컬러공간에서 HOG(Histogram of Oriented Gradient) 특징을 추출한 후 SVM 분류기를 이용하여 신호등을 인식하였다. Weber 외[13]는 신호등 영역의 검출 및 식별을 모두 심층 신경망을 적용하는 방식을 제안하였으며, Behrendt 외[14]는 심층 신경망을 통해 신호등을 검출한 후 연속된 프레임에서 신호등을 추적하여 인식률을 개선하였다.

3. 신호등 점등영역 검출

신호등은 검정색 함체에 신호를 나타내기 위한 점멸등이 배치되어 있다. 3색등은 적색, 황색, 녹색의 점멸등이 배치되어 있고, 4색등은 여기에 좌회전 표시등이 추가된다. 점멸등은 영상 내 다른 영역에 비하여 밝게 나타나고, 신호등 함체는 검정색을 띤 직사각형의 형태를 나타낸다. 이러한 특징은 신호등 검출을 위한 중요한 실마리가 된다. 신호등 검출을 위하여 점등영역과 신호등 함체를 모두 탐색하는 기존 연구들[1-3, 8]과는 달리 본 연구에서는 점등영역만을 탐색한다. 신호등의 점등영역은 일조량이나 촬영 각도에 따라 다르게 나타나기 때문에 기존 연구들에서는 점등영역의 검출 실패를 줄이기 위하여 색 구분 임계값의 범위를 여유 있게 설정하고 있다. 그러나 색 구분을 위하여 정의한 범위가 연구들 마다 상당한 차이를 보이고 있고, 점등영역에 과포화가 발생하는 경우 색 구분이 곤란하다.

본 연구에서는 신호등 점등영역 검출 단계에서 색상별 범위를 정의하지 않고 명도와 채도를 이용하는 접근 방식[9]을 채택하였다. 신호등 점등영역은 조도

에 따라 색상의 변화가 크게 나타나지만, 명도와 채도는 상대적으로 안정적인 상태를 나타낸다. 일반적인 경우 점등영역은 명도와 채도가 높게 나타나며 점등영역에 과포화가 발생하는 예외적인 경우에는 채도는 낮지만 명도는 높게 나타난다. 따라서 이러한 두 가지 경우를 고려하여 신호등 점등영역 검출을 위한 특징 영상을 생성한다. Fig. 1은 점등영역 영상 I_F 를 생성하는 전반적인 과정을 나타낸 것이다.

3.1 점등 후보영역 추출

RGB 컬러영상을 HSV 컬러영상으로 변환하여 대응되는 화소의 채도(S)와 명도(V) 값을 곱하여 실수 영상 $I_{S \times V}$ 을 생성한다. 이 영상은 신호등의 점등영역이 충분히 밝게 나타나지 않아도 채도 값에 의해 보상되며, 채도가 다소 낮게 나타나더라도 명도 값에 의해 보상되는 효과를 갖는다. 영상 $I_{S \times V}$ 의 화소 값들을 0부터 255까지의 범위로 정규화한 영상 I_N 을 임계값 th_1 로 분할하여 점등영역이 전경으로 나타나는 이진영상 B_1 을 구한다. 식 (1)에서 th_1 은 120으로 설정하였다. Fig. 2-(a)는 블랙박스 입력영상으로, 적색 사각형으로 표시된 관심영역(ROI, region of interest)을 대상으로 신호등 점등영역을 추출한다. Fig. 2-(b)에 제시된 ROI 영상에서 왼쪽 신호등에서 과포화로 인해 채도가 매우 낮게 나타나 Fig. 2-(c)의 이진영상 B_1 에 점등영역이 추출되지 않았다.

$$B_1(x, y) = \begin{cases} 1, & I_N(x, y) \geq th_1 \\ 0, & otherwise \end{cases} \quad (1)$$

신호등 점등영역에 과포화가 발생하면 채도는 매우 낮게 나타나지만 명도는 높게 나타나므로, 이 경우에는 밝기 값만으로 신호등 점등영역을 탐색한다. 신호등 점등영역의 채널별 명도 특성을 이용하여 식 (2)에 의해 이진영상 B_2 를 생성한다. 식 (2)에서 I_R, I_G, I_B 는 RGB 컬러영상을 R, G, B 채널별로 분리한 명도영상으로, 본 연구에서 th_2 는 150으로 설정하였

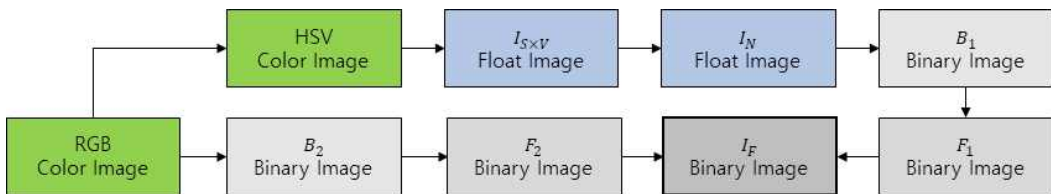


Fig. 1. Overall process for creating an image of lighting regions.

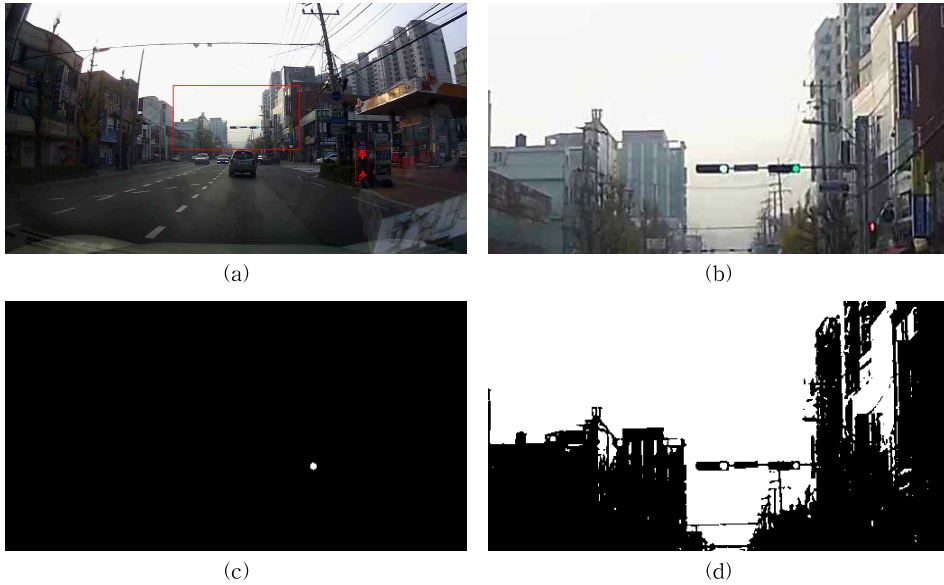


Fig. 2. (a) input image, (b) ROI image (c) binary image B_1 , (d) binary image B_2 .

다. 따라서 G채널 화소의 명도 값이 150이상이고, R 채널 또는 B채널 화소의 명도 값이 150이상인 화소들이 이진영상 B_2 의 전경이 된다.

$$B_2(x, y) = \begin{cases} 1, & I_G(x, y) \geq th_2 \text{ AND } (I_R(x, y) \geq th_2 \text{ OR } I_B(x, y) \geq th_2) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

3.2 점등영역 영상 생성

이진영상 B_1 과 B_2 에는 신호등 점등영역 외에 여러 다른 영역들이 포함되어 나타날 수 있다. 영상 B_1 과 B_2 에 존재하는 거짓긍정(FP, false positive) 영역을 신호등 인식단계 이전에 미리 제거한다면 전체적인 처리 속도를 개선할 수 있으므로, 본 연구에서는 FP 영역을 제거하기 위하여 형태학 열림 연산을 이용하여 잡영을 제거하고, 연결요소에 대하여 윤곽선을 추

출한 후 크기와 형태 분석을 수행한다. 자동차에 부착된 카메라와 신호등과의 거리가 가까워질수록 신호등의 위치는 영상의 위쪽으로 이동하고, 신호등의 크기는 커지게 된다. 본 연구에서는 윤곽선의 반경이 3픽셀 미만이거나 20픽셀을 초과하는 윤곽선에 해당하는 영역들을 FP영역으로 판단하여 제거하였다. 이진영상 B_1 과 B_2 에 존재하는 연결요소들을 영역 크기에 의해 필터링하여 생성한 영상 F_1 과 F_2 를 식 (3)에 의해 병합하여 점등영역 영상 I_F 를 생성한다. Fig. 3-(a)는 Fig. 2-(d)로부터 연결요소의 크기를 기준으로 필터링한 결과 영상이고, Fig. 3-(b)는 필터링을 거친 두 영상을 병합한 점등영역 영상을 예시한 것이다.

$$I_F(x, y) = \max(F_1(x, y), F_2(x, y)) \quad (3)$$

신호등의 점등영역은 원형이므로 식 (4)에 의해



Fig. 3. (a) size filtered image F_2 , (b) merged image I_F .

윤곽선의 원형도(roundness)를 계산하여 점등영역 영상 I_p 에 남아 있는 FP영역을 제거한다. a 는 면적을 의미하고, d 는 직경을 나타내는 것으로 ρ 가 클수록 윤곽선이 원형에 가깝다는 것을 의미한다[15]. 본 연구에서는 ρ 가 5.0 보다 작은 영역을 제거하였다.

$$\rho = \frac{4\pi a}{d^2} \quad (4)$$

4. 신호등 인식

점등영역 영상에 나타난 연결요소들이 신호등의 일부인지, 만약 그렇다면 신호등은 어떤 신호를 나타내는지를 식별하기 위하여 심층 합성곱 신경망(deep convolutional neural network, DCNN)을 도입하였다. 본 연구에서 신호등 인식은 적색(적색등), 통과(녹색등), 진입시 통과(황색등)의 3가지 신호를 식별 대상으로 한다.

4.1 DCNN 입력 데이터

점등영역 영상에 나타난 연결요소들을 기반으로 신호등을 검출하게 되는데, 점등영역의 색상 정보만으로는 신호등 여부를 정확하게 판단하기 어렵고, 점등영역에 과포화가 발생할 경우 점등된 부분의 색상을 구별할 수 없다. 신호등은 적색등, 황색등, 녹색등의 순서로 좌에서 우로 배치되어 있으므로, 점등영역의 위치가 신호를 식별하는데 중요한 정보가 된다. 따라서 본 연구에서는 점등영역을 확장하여 DCNN 입력 데이터를 생성하였다. Table 1은 입력 데이터에 포함된 각 클래스별 데이터의 개수를 나타낸 것이다. Fig. 4는 입력 데이터를 예시한 것이다. 적색등과 녹

색등은 점등영역을 중심으로 각각 우측과 좌측으로 점등영역 크기를 확장하였고, 황색등은 점등영역을 중심으로 좌우로 확장하여 훈련 데이터를 생성하였다. 점등영역 영상에는 신호등 점등영역이 아닌 다른 영역, 즉 FP(false positive)영역이 남아 있을 수 있다. 따라서 FP영역을 식별하기 위해 비신호등, 즉 배경에 대한 입력 데이터도 요구된다. 배경 데이터는 신호등이 존재하지 않는 영상으로부터 랜덤으로 위치와 크기를 선택하여 생성하였다.

4.2 DCNN 모델 및 학습

신호등을 인식하기 위하여 본 연구에서 사용한 DCNN 모델의 구조는 Table 2와 같다. 입력 데이터는 크기가 32×8로 정규화된 RGB 컬러영상이며, DCNN은 3개의 합성곱 신경망과 2개의 완전연결 신경망으로 구성되어 있다. 커널의 크기는 (3, 3)이고, 활성화 함수는 최종 단계에서 softmax를, 그 이전 단계에서는 Relu를 사용하였다. 그리고 신경망의 과적합(over fitting) 문제를 해결하기 위하여 드롭아웃(dropout)을 적용하였다.

DCNN 모델의 학습과 평가를 위하여 입력 데이터의 75%는 훈련 데이터, 25%는 테스트 데이터로 분리시켰다. 학습율은 0.001, 배치 크기는 32, 옵티마이저는 adam, 전체 훈련 데이터에 대한 반복 학습의 횟수는 50으로 설정하여 DCNN 모델을 학습시킨 후 DCNN 모델을 평가한 결과 정확도(accuracy) 99.91%를 얻었다.

4.3 신호등 식별

제안된 DCNN 분류기의 입력 영상은 점등영역의

Table 1. Each class size of the training data

	Red Light	Yellow Light	Green Light	Background
No. of data	68	60	222	1,850

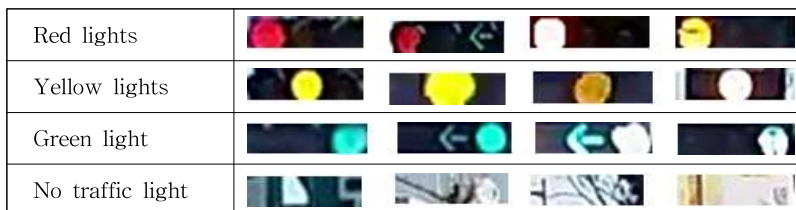


Fig. 4. Samples of the training data.

Table 2. Architecture of the DCNN classifier

Input	32×8×3
Convolution	filter: 32, kernel: (3, 3), activation: Relu
MaxPooling	kernel: (2, 2)
Dropout	rate: 0.25
Convolution	filter: 64, kernel: (3, 3), activation: Relu
Convolution	filter: 64, kernel: (3, 3)
MaxPooling	kernel: (2, 2)
Dropout	rate: 0.25
Fully connected	units: 512, activation: Relu
Dropout	rate: 0.5
Fully connected	units: 4, activation: softmax

색상과 위치 정보를 얻기 위하여 점등영역을 확장한 영상이 입력이 된다. 점등영역의 색상에 따라 점등영역을 좌, 우, 또는 좌우로 확장하여야 하는데, DCNN 분류기의 입력 단계에서는 점등영역이 어떤 색상인지 미리 알 수 없다. 그러므로 점등영역 영상내의 연결요소에 대하여 최소의 접사각형을 구한 후 좌측으로 확장한 영상, 좌우로 확장한 영상, 우측으로 확장한 영상을 모두 DCNN 분류기의 입력 영상으로 사용한다. 즉, 한 개의 점등영역에 대하여 총 3개의 점등영역 확장 영상이 만들어져 DCNN 분류기로 입력된다. Fig. 5는 한 개의 점등영역으로부터 생성된 3개의 영상 $I_1 \sim I_3$ 을 예시한 것으로, 제시된 사각형은 점등영역이 어떤 방식으로 확장되는지 이해를 돕기 위하여 영상 위에 중첩하여 표시한 것이다.

한 개의 점등영역에 대하여 3개의 영상이 DCNN 분류기에 입력이 되므로, 3개의 인식 결과가 출력된다. 3개 영상 각각에 대한 인식 결과를 종합하여 점등영역에 대한 최종 인식 결과를 도출하는 알고리즘은 아래와 같다. 3개 중 한 개만 신호등으로 인식되고, 나머지 두 개는 배경으로 식별되면 인식된 신호등의

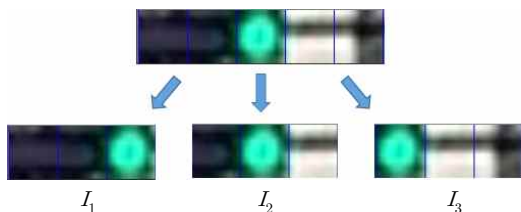


Fig. 5. Input images of the DCNN classifier.

결과를 출력한다. 예를 들어 Fig. 5의 경우 왼쪽 첫 번째 영상인 I_1 이 녹색등으로 인식되고, 나머지 두 개는 배경으로 인식되면 해당 점등영역은 녹색 신호등으로 인식된다. 점등영역이 황색등인 경우, 첫 번째 입력 영상 I_1 은 배경으로 인식되고, 2~3번째 입력 영상 I_2 와 I_3 는 각각 황색등과 적색등으로 인식될 수 있다. 따라서 이와 같은 경우는 황색등으로 인식 결과를 출력한다. 이외의 나머지 경우는 모두 배경으로 판정한다.

알고리즘: 3개 영상에 대한 인식 결과로부터 최종 인식결과를 도출

```

 $N_B$  = 배경으로 인식된 영상의 개수
if (  $N_B == 2$  )
    최종 인식 결과 = 나머지 1개 영상에 대한 인식 결과
else if (  $N_B == 1$  ) and ( 영상  $I_2 ==$  황색등 and 영상  $I_3 ==$ 
    녹색등 )
    최종 인식 결과 = 황색등
else
    최종 인식 결과 = 배경
    
```

5. 실험 결과 및 분석

5.1 실험 데이터

실험을 통해 제안한 신호등 인식 방법의 타당성을 확인하기 위하여 훈련 데이터와는 별도로 주간에 도심 지역에서 촬영한 총 12개의 동영상상을 테스트 데이터로 이용하였다. 동영상상은 30fps로 촬영되었으며, 각 프레임은 1,920×1,080의 해상도를 갖는다. 동영상내의 모든 프레임을 처리하려면 처리 시간이 길어지므로, 본 실험에서는 5개 프레임 간격으로 영상을 처리하였다. Table 3은 실험 데이터에서 영상 중앙부의 신호등 탐지 영역에 존재하는 신호등에 대한 검증자료(ground truth)이다. 신호등 탐지 영역에 존재하는 신호등의 최소 크기는 반경이 3픽셀이다. 동영상 1번~6번은 겨울에 촬영한 것이고, 동영상 7번~12번은 여름에 촬영한 것이다.

5.2 성능 측정

제안된 신호등 인식 방법에 대한 성능은 프레임 단위로 측정하였다. 신호등 인식은 신호등 검출과 식

Table 3. Ground truth of test data

Video Clip	# of total frames	# of processed frames			
		Red	Yellow	Green	None
01~06	3,193	7	27	242	360
07~12	2,374	80	0	132	260
Total	5,567	87	27	374	620

별의 두 가지로 나누어 살펴볼 수 있다. 신호등 검출에 대한 성능은 정밀도(precision)와 재현율(recall)로 나타낼 수 있다. 정밀도는 신호등이 존재한다고 판단한 프레임 중에서 신호등이 실제 존재하는 프레임의 비율을 나타낸 것이다. 식 (5)에서 TP는 참 긍정(true positive)으로 신호등이 존재할 때 신호등을 검출한 프레임의 개수, FP는 긍정 오류(false positive)로 신호등이 없는데 신호등을 검출한 프레임의 개수를 나타낸다. 재현율은 실제로 신호등이 존재할 때 신호등이 존재한다고 판단한 프레임의 비율을 나타낸 것이다. 식 (6)에서 FN은 부정 오류(false negative)로 신호등이 존재하는데 신호등을 검출하지 못한 프레임의 개수를 의미한다. Table 4에 제시된 바와 같이 제안하는 방법은 정밀도 93.9%, 재현율 91.6%의 신호등 검출 성능을 보였다.

$$Precision(\%) = \frac{TP}{TP+FP} \times 100 \quad (5)$$

$$Recall(\%) = \frac{TP}{TP+FN} \times 100 \quad (6)$$

신호등 식별에 대한 성능은 식 (7)의 정확도(accuracy)로 측정하였다. 정확도는 신호등의 상태를 올바르게 식별한 프레임의 비율을 나타낸 것이다. TC

(true class)와 FC(false class)는 신호등이 존재할 때 신호를 올바르게 식별한 프레임이 개수와 그릇되게 식별한 프레임의 개수이다. TN은 참 부정(true negative)으로 신호등이 없을 때 신호등이 없다고 판단한 프레임의 개수를 나타낸다. 제안한 방법은 Table 5에 제시된 바와 같이 89.4%의 정확도를 보였다.

$$Accuracy(\%) = \frac{TC+TN}{TC+TN+FP+FC} \times 100 \quad (7)$$

DCNN 학습에 사용된 데이터의 개수가 비교적 작은 편이기 때문에 케라스(Keras)에서 제공하는 Image DataGenerator 클래스를 이용하여 학습데이터를 확장하였을 때, 인식 성능이 얼마나 개선되는지 실험하였다. 학습 데이터를 확장하기 위하여 회전 및 이동 변환을 수행하였다. 회전 변환은 ±10도, 이동 변환은 좌우, 상하 각각 10% 범위내로 하였다. 적색, 녹색, 황색 신호등 데이터는 각 영상 당 10개, 비신호등 데이터는 각 영상 당 5개의 영상을 생성하였다. 이렇게 하여 748개의 적색 신호등, 660개 황색 신호등, 2,442개의 녹색 신호등, 그리고 11,100개 배경 영상을 학습데이터로 사용하였다. Table 6은 확장된 학습데이터로 DCNN을 학습시켜 신호등 인식 성능을 측정한 결과이다. 학습데이터를 확장하기 전에 비하여 재현

Table 4. Performance of traffic light detection

Video Clip	TP	FP	TN	FN	Precision	Recall
01~06	255	12	348	21	95.5%	92.4%
07~12	192	17	243	20	91.9%	90.6%
Total	447	29	591	41	93.9%	91.6%

Table 5. Performance of traffic light recognition

Video Clip	Traffic light exist		No traffic light		Accuracy
	TC	FC	TN	FP	
01~06	224	52	348	12	89.9%
07~12	175	37	243	17	88.6%
Total	399	89	591	29	89.4%

Table 6. Performance of the DCNN using augmented data

Video Clip	traffic light detection		recognition
	Precision	Recall	Accuracy
01~06	85.9%	92.8%	88.4%
07~12	87.6%	94.5%	86.9%
Total	86.6%	93.5%	87.7%

율은 약간 향상되었으나, 정밀도와 인식의 정확도는 오히려 성능이 저하된 것으로 나타났다. 점등 영역을 토대로 추출한 신호등 영상을 신호등이 아니라고 판단하는 부정오류(FN)는 감소하였으나, 배경 영상을 신호등으로 판단하는 거짓긍정(FP)이 늘어났기 때문이다. 이러한 결과는 회전 변형과 좌우상하 이동에 의한 변형이 신호등 영상에 나타날 수 있는 다양한 특성을 표현하기 보다는 오히려 신호등 영상에 나타나는 특성을 왜곡시켰기 때문인 것으로 판단된다.

Table 7은 제안된 방법과 기존 연구[11, 13]의 실험 결과들을 정리한 것이다. 신호등 인식 성능은 동영상을 획득할 당시의 계절, 날씨, 영상의 해상도에 따라서 차이를 보이며, 신호등 검출 및 인식을 위한 카메라와 신호등간의 최대 거리에 따라라도 성능에 차이를 보인다. 예를 들어 최대 거리가 50 m라면 동영상의 각 프레임에서 카메라와 신호등간의 거리가 50 m이내에 존재하는 신호등만을 인식 대상으로 함을 의미한다. 즉 최대 거리를 벗어나 너무 작게 나타나는 신호등은 인식 대상에서 제외하는 것이다. 일반적으로 신호등이 작게 나타날수록 인식 오류가 발생하므로 최대 거리가 멀수록 인식 성능은 저하된다. 실험 데이터가 동일하지 않은 상태에서 기존 연구 결과들과 직접적인 비교는 곤란하다. 그렇지만 제시된 실험 조건을 감안하여 비교해 본다면 제안한 방법의 성능을 개략적으로 가늠해 볼 수 있다. Table 7에서 제안된 방법, Jang 외[11], Weber 외[13]의 연구들

은 카메라로부터 획득된 영상만을 이용한 결과이고, Jang 외[11]⁺는 영상은 물론 위치기반 시스템과 디지털 지도를 부가적으로 활용한 결과이다. Weber 외[13]의 연구는 최대거리가 명시되지 않았으나 해상도가 낮은 경우 성능이 저하됨을 알 수 있다. 입력 영상 이외에 추가적인 정보 없이 신호등을 인식한 기존 연구들과 비교해 보았을 때, 각 프레임 당 평균 처리 속도는 물론 신호등 검출 및 인식 성능 면에서도 제안된 방법이 효과적임을 확인할 수 있었다.

6. 결 론

색상은 신호등 검출 및 식별을 위한 중요한 특징이지만, 야외에서 획득되는 영상의 특성상 조도 변화에 의하여 색상이 민감하게 변화한다. 특히 신호등 점등 영역에 과포화가 발생하면 적색등, 녹색등, 황색등이 모두 흰색으로 나타난다. 이러한 문제를 해결하기 위하여 본 연구에서는 조도 변화에 강인한 신호등 점등 영역 추출 방법을 도입하였으며, 검출된 영역을 확장한 후 심층 합성곱 신경망을 통해 신호등을 인식하는 방법을 제안하였다. 기존 연구에서는 대개 신호등 합체를 검출하여 신경망 입력으로 사용하는 데 비하여, 본 연구는 점등 영역만을 검출한 후 점등 영역을 좌측, 우측, 좌우로 확장한 영상을 신경망 입력으로 사용하였다. 이 방식은 신호등 색상의 식별을 신호등 인식 단계로 늦춤으로써 신호등 검출 단계에

Table 7. Performance comparison of traffic light recognition

Method	Max distance	Image Resolution	Processing Time	Detection		Recognition
				Precision	Recall	Accuracy
Jang et al.[11]	80 m	2,040×1,086	178.8 ms	86.9%	78.4%	81.3%
Jang et al.[11] ⁺	80 m	2,040×1,086	11.2 ms	99.7%	92.8%	95.5%
Weber et al.[13]	NA	640×480	28.8 ms	90.6%	96.3%	88.2%
		1,280×960	80.0 ms	95.8%	91.5%	96.5%
Proposed	70 m	1,920×1,080	32.3 ms	93.9%	91.6%	89.4%
	50 m			95.5%	99.6%	98.0%

서 색상 변화에 의한 잘못된 판단을 유보하는 이점이 있다. 도심에서 중간에 획득한 12개의 동영상에 대하여 신호등 인식 실험을 수행한 결과, 제안하는 방법은 신호등 검출의 정밀도는 93.9%, 재현율은 91.6%였으며, 신호등 인식의 정확도는 89.4%를 보여 제안하는 방법이 효과적임을 확인할 수 있었다. 그러나 과포화에 의한 흰색을 점등 영역으로 포함시키면서 배경의 일부가 점등 영역으로 오인식되는 경우가 발생하기도 하였다. 따라서 신호등 검출 및 식별 단계에서 점등 영역 주변의 텍스처 정보를 추가적으로 도입하여 성능을 개선할 계획이다.

REFERENCE

- [1] Z. Cai, Y. Li, and M. Gu, "Real-time Recognition System of Traffic Light in Urban Environment," *Proceedings of the 2012 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pp. 1-6, 2012.
- [2] P. Jo and J. Lee, "Traffic Light Detection Using Morphometric Characteristics and Location Information in Consecutive Images," *Journal of Institute of Control, Robotics and Systems*, Vol. 21, No. 12, pp. 1122-1129, 2015.
- [3] J. Kim and J. Kim, "Performance Improvement of Traffic Signal Lights Recognition Based on Adaptive Morphological Analysis," *Journal of the Korea Institute of Information and Communication Engineering*, Vol. 19, No. 9, pp. 2129-2137, 2015.
- [4] J. Jeong and D. Rho, "Real Time Detection and Recognition of Traffic Lights Using Component Subtraction and Detection Masks," *Journal of the Institute of Electronics Engineers of Korea*, Vol. 43, No. 2, pp. 65-72, 2006.
- [5] T. Gevers and A. Smeulders, "Color-based Object Recognition," *Pattern Recognition*, Vol. 32, No. 3, pp. 453-464, 1999.
- [6] M. Omachi and S. Omachi, "Traffic Light Detection with Color and Edge Information," *Proceedings of the IEEE International Conference on Computer Science and Information Technology*, pp. 284-287, 2009.
- [7] M. D.-Cabrera, P. Cerri, and J. S.-Medina, "Suspended Traffic Lights Detection and Distance Estimation Using Color Features," *Proceedings of the IEEE International Conference on Intelligent Transportation Systems*, pp. 1315-1320, 2012.
- [8] H. Moizumi, Y. Sugaya, M. Omachi, and S. Omachi, "Traffic Light Detection Considering Color Saturation Using In-Vehicle Stereo Camera," *Journal of Information Processing*, Vol. 24, No. 2, pp. 349-357, 2016.
- [9] M. Kim, "Detection of a Light Region Based on Intensity and Saturation and Traffic Light Discrimination by Model Verification," *Journal of Korea Multimedia Society*, Vol. 20, No. 11, pp. 1729-1740, 2017.
- [10] Z. Ozcelik, C. Tastimur, M. Karakose, and E. Akin, "A Vision Based Traffic Light Detection and Recognition Approach for Intelligent Vehicles," *Proceedings of the International Conference on Computer Science and Engineering*, pp. 424-429, 2017.
- [11] C. Jang, S. Cho, S. Jeong, J. Suhr, and H. Jung, "Traffic Light Recognition Exploiting Map and Localization at Every Stage," *Expert Systems with Applications*, Vol. 88, pp. 290-304, 2017.
- [12] X. Zhou, J. Yuan, and H. Liu, "Real-Time Traffic Light Recognition Based on C-HOG Features," *Computing and Informatics*, Vol. 36, No. 4, pp. 793-814, 2017.
- [13] M. Weber, P. Wolf, and J. M. Zollner, "Deep TLR: A Single Deep Convolutional Network for Detection and Classification of Traffic Lights," *Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 342-348, 2016.
- [14] K. Behrendt, L. Novak, and R. Botros, "A Deep Learning Approach to Traffic Lights: Detection, Tracking, and Classification," *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1370-1377, 2017.

- [15] V. John, K. Yoneda, Z. Liu, and S. Mita, "Saliency Map Generation by the Convolutional Neural Network for Real-Time Traffic Light Detection Using Template Matching," *IEEE Transactions on Computational Imaging*, Vol. 1, No. 3, pp. 159-173, 2015.



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