



LGP-FL과 해마 구조를 이용한 H-CNN 기반 보행자 검출에 대한 연구

박수빈*, 강대성**

A Study on H-CNN Based Pedestrian Detection Using LGP-FL and Hippocampal Structure

Su-Bin Park*, Dae-Seong Kang**

이 논문은 2017 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2017R1D1A1B04030870)

요약

최근 자율 주행 자동차에 대한 연구가 활발하다. 자율 주행 자동차는 보행자 검출 및 인식 기술이 중요하다. 최근에 주로 사용되는 CNN(Convolutional Neural Network)을 이용한 보행자 검출은 대체로 좋은 성능을 보이나 영상의 환경에 따른 성능 저하가 있다. 본 논문에서는 LGP-FL(Local Gradient Pattern-Feature Layer)을 추가한 CNN Network를 기반으로 해마 신경망의 장기 기억 구조를 적용한 보행자 검출 시스템을 제안한다. 먼저 입력 이미지를 227x227의 크기로 변경한다. 그 후 총 5개 층의 Convolution layer를 거쳐 특징을 추출한다. 그 과정에서 추가되는 LGP-FL에서는 LGP 특징 패턴을 추출하여 출현 빈도수가 높은 패턴을 장기 기억 장치에 저장한다. 이후 검출 과정에서 밝기 및 색상 변화에 강한 LGP 특징 패턴 정보를 이용해 검출함으로써 보다 정확하게 보행자를 검출할 수 있다. 기존의 방법들과 제안하는 기법의 비교를 통해 약 1~4%의 검출률 증가를 확인하였다.

Abstract

Recently, autonomous vehicles have been actively studied. Pedestrian detection and recognition technology is important in autonomous vehicles. Pedestrian detection using CNN(Convolutional Neural Network), which is mainly used recently, generally shows good performance, but there is a performance degradation depending on the environment of the image. In this paper, we propose a pedestrian detection system applying long-term memory structure of hippocampal neural network based on CNN network with LGP-FL (Local Gradient Pattern-Feature Layer) added. First, change the input image to a size of 227x227. Then, the feature is extracted through a total of 5 layers of convolution layer. In the process, LGP-FL adds the LGP feature pattern and stores the high-frequency pattern in the long-term memory. In the detection process, it is possible to detect the pedestrian more accurately by detecting using the LGP feature pattern information robust to brightness and color change. A comparison of the existing methods and the proposed method confirmed the increase of detection rate of about 1~4%.

Keywords

pedestrian detection, object detection, LGP, CNN, hippocampal structure

* 동아대학교 전자공학부
 - ORCID: <https://orcid.org/0000-0001-6346-8960>
 ** 동아대학교 전자공학과 교수(교신저자)
 - ORCID: <https://orcid.org/0000-0003-0186-2430>

• Received: Aug. 13, 2018, Revised: Oct. 22, 2018, Accepted: Oct. 25, 2018
 • Corresponding Author: Dae-Seong Kang
 Dept. of Dong-A University, 37 NaKdong-Daero 550, beon-gil saha-gu, Busan, Korea,
 Tel.: +82-51-200-7710, Email: dskang@dau.ac.kr

1. 서 론

최근 자율 주행 자동차에 대한 연구가 활발하다. 자율 주행 자동차는 최종적으로 인간의 제어 없이도 완벽하게 도로를 주행하는 것이 목표이다. 자율 주행 자동차의 주요 기술 중 하나에는 환경인식이 있는데, 그중 차량 주변의 보행자의 검출 및 인식은 안전과 직결되어 있기 때문에 정확성이 매우 중요하다. 대부분의 검출 기술들은 입력 영상으로부터 보행자의 특징을 추출해서 학습된 분류기를 통해 분류하는 방법을 이용한다. 특징 추출 방법에는 영상의 밝기 차이를 이용하는 Haar-like, LBP(Local Binary Pattern) 등이 있다[1][2]. 그리고 영상 내 기울기의 변화를 이용하는 HOG(Histogram of Oriented Gradients) 특징과 여러 크기의 HOG 특징을 조합하여 사용하는 Cascade HOG 등이 있다[3][4]. 위에 언급한 특징 추출 방법들은 Hand-crafted 방식으로서 설계자가 직접 특징 값에 대해서 설계하기 때문에 성능의 한계가 있었다.

최근에는 딥 러닝(Deep Learning)을 이용한 특징 추출 및 학습 방법이 주로 사용된다. 딥 러닝은 기계 학습의 한 분야로서 기존의 인공 신경망의 층을 깊게 쌓은 신경망을 뜻한다. 하드웨어의 발전으로 방대한 계산량을 감당할 수 있게 되었고 때문에 최근 다시 활발하게 연구 되고 있는 분야이다. 영상 처리에서는 CNN(Convolutional Neural Network) 신경망 구조가 대부분 사용된다. CNN은 컨볼루션(Convolution) 연산을 기반으로 특징을 추출하는 방법으로서 대표적인 CNN 구조로는 AlexNet[5], VGGNet[6], ResNet[7] 등이 있다.

본 논문에서는 AlexNet 구조에 해마 구조를 적용한 H-CNN(Hippocampal-CNN) 기반 보행자 검출 시스템을 제안한다. 기존의 AlexNet을 이용한 보행자 검출의 경우 과거의 Hand-crafted 방식을 사용한 방법들보다 우수한 성능을 보인다. 그러나 입력 영상의 환경에 따라서 성능이 저하되는 경우가 있다[8]. 따라서 본 논문에서는 AlexNet을 기반으로 해마 구조의 장기 기억 장치의 구현을 통해 보행자의 특징 패턴을 미리 기억하고 검출에 사용하여 보다 정확한 검출을 가능하게 한다. 보행자의 특징 패턴 추출에는 LGP(Local Gradient Pattern) 특징추출 방법을

사용하였다[9]. LGP 특징은 밝기 차의 평균값을 이용하기 때문에 주변 환경의 변화나 밝기의 변화 등에 강인한 장점이 있다. 기존의 연구로는 CNN의 전처리 과정에서 LGP 특징을 이용하여 학습 데이터를 검출 목적에 맞게 전 처리하여 학습에 사용한 방법이 있다[10]. 언급한 방법은 검출률은 우수하지만 처리 과정이 전처리 과정과 학습 과정 2단계로 구성되기 때문에 전체적인 학습 속도가 오래 걸리는 단점이 있었다. CNN은 컨볼루션 필터를 이용해서 특징을 스스로 학습 하는 능력이 있기 때문에 전처리 과정 없이도 특징 추출 능력은 충분히 우수하다. 따라서 본 논문에서는 전처리 과정은 없애고 CNN 내부에 LGP-FL(Local Gradient Pattern-Feature Layer)을 추가하였다. LGP-FL에서는 LGP 특징 패턴을 추출하여 해마 구조의 장기 기억 장치에 출현 빈도순으로 저장하는 역할을 한다. 그 후 검출 단계에서 빈도순으로 저장된 LGP 특징 패턴 정보를 이용하여 보다 정확하게 보행자를 검출할 수 있다. 실험에는 INRIA 보행자 Dataset을 이용하여 기존의 방법들과 제안하는 기법의 비교를 통해 성능을 평가하였다.

II. 관련 이론

2.1 HOG

HOG는 입력 영상에서 특정 영역에 대한 밝기의 기울기를 계산하여 그 값들을 특징으로 사용하는 특징 추출 방법이다. HOG 특징을 구하는 방법은 입력 영상을 Sliding window 방식으로 탐색하면서 미리 정해진 크기의 셀로 분할하고 각각의 셀들의 방향 값들을 히스토그램으로 나열한다[3].

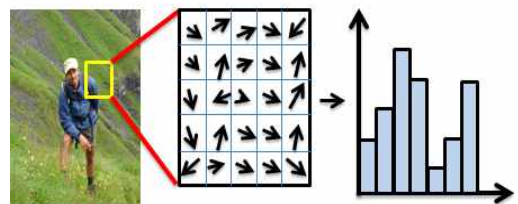


그림 1. HOG 특징 추출 과정
Fig. 1. HOG Feature extraction process

HOG 특징은 조명의 변화에 강하고 지역적 영역의 여러 가지 변화에도 강한 특징이 있다. HOG 특징 추출 과정은 그림 1과 같이 나타낼 수 있다.

2.2 LGP

LGP는 기존 LBP[2]의 조명의 변화, 주변 배경의 변화에 민감한 단점을 개선하여 변화에 대한 민감도를 최소화하기 위한 방법이다. 다음과 같은 수식으로 나타낼 수 있다.

$$LGP(x_a, y_a) = \sum_{n=0}^7 s(|g_n - g_c| - \Delta M) 2^n \quad (1)$$

$$\Delta M = \frac{\sum_{n=0}^7 |g_n - g_c|}{8}$$

$$s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

식 (1)에서 x_a, y_a 는 중앙 픽셀 위치, g_n 은 이웃 픽셀 값, g_c 는 중앙 픽셀 값을 나타낸다. 그림 2를 보면 중앙 픽셀 값과 이웃 픽셀 값들의 차이 값 즉 기울기를 모두 더한 다음 그 값의 평균을 구해 중앙 픽셀 값과 비교하여 1과 0으로 2진화하여 총 8비트의 값을 얻어 256가지의 패턴으로 변환시키게 된다.

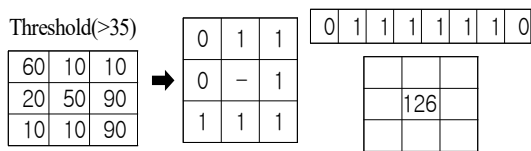


그림 2. 3x3 LGP
Fig. 2. 3x3 local gradient pattern

기존의 LBP는 이웃 픽셀 값과 중앙 픽셀 값의 크기 자체를 비교하기 때문에 작은 변화에도 값이 변할 수 있어 작은 변화에도 민감한 경우가 생기는데, LGP는 밝기 차이들의 평균값을 사용하기 때문에 작은 변화에도 강인한 이점이 있다[9].

2.3 CNN

CNN은 크기 조절이 가능한 여러 개의 필터를 사용해 입력 영상으로부터 의미 있는 특징을 추출하는 과정을 가진 인공신경망의 한 종류이다. 최초의 CNN 구조는 1998년 Y.Lecun의 LeNet[11]으로 당시에는 깊은 신경망에 따른 많은 연산량, 학습 데이터의 부족 등의 문제로 한계를 겪었으나 최근 GPU를 기반으로 한 하드웨어의 발달과 방대한 공개 Dataset들에 의해서 활발히 연구가 이루어지고 있다.

CNN의 구조는 크게 3단계에 걸쳐 구성된다. 첫 번째는 Convolution layer로 입력 영상에 필터를 통과시켜 Sliding-window 방식으로 연산하며 의미 있는 특징들을 추출해낸다. 두 번째는 Pooling layer로 Convolution 연산을 거쳐 추출된 특징들의 크기를 줄이는 Subsampling 단계이다. 대부분의 이미지의 인접한 픽셀들은 값이 비슷한 경우가 많기 때문에 정보를 보존하면서 크기를 줄여주는 역할을 한다. Pooling은 Max pooling과 Average pooling 방법을 주로 사용한다. 마지막은 Fully-connected layer로 과거 신경망의 연산 방식처럼 모든 뉴런들을 연결하는 부분이다. 컨볼루션 연산을 통해 최종적으로 구해진 잘 압축된 특징 정보들의 Raw 데이터 값들을 마지막에 모두 연결시켜 학습 한다. 그림 3에서 CNN의 구조를 간단하게 나타내었다.

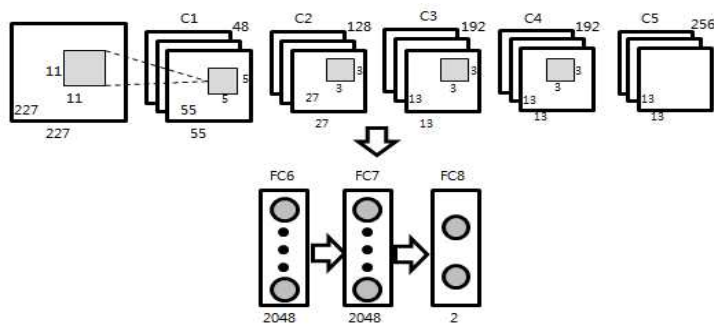


그림 3. CNN의 구조
Fig. 3. Structure of CNN

2.4 해마 신경망

해마 신경망(Hippocampal Neural Network)은 인간의 뇌에서 기억을 담당하는 해마의 구조를 공학적으로 모델링 한 신경망이다. 해마는 시각을 통해 정보가 들어오면 그 정보에 대한 기억을 하는 역할을 한다.

기억은 크게 단기 기억과 장기 기억으로 나뉘는데 시각을 통해 정보가 들어오면 그 정보는 단기 기억 장소로 입력되고 작게는 몇 분에서 길게는 며칠간 저장되었다가 삭제된다. 그 과정에서 만약 같은 정보가 반복되면 그 정보는 단기 기억 장소에서 장기 기억 장소로 옮겨지게 되고 그 기억은 반영구적이다. 신경망에서 특징을 학습할 때 출현이 잦은 특징 패턴이 존재한다면 그 특징 패턴을 장기 기억 장치에 저장해두고 사용할 수 있다[12].

그림 4는 해마 구조의 모델링을 나타낸다. 각 블록의 기능으로는 먼저 내후각뇌피질(EC)에서 신경망 모델의 입력과 출력을 나타내고 그 후 치아이랑(DG)에서 입력으로 들어온 정보의 특징을 단순화시키는 역할을 한다. 본 논문에서는 LGP 특징 패턴을 이용하여 단순화 시킨다. 단순화된 특징은 CA3 영역으로 저장되어 삭제 혹은 기억되는데 특징 패턴의 정보가 반복적으로 학습되는 경우 CA1로 저장되는 시스템으로 이루어져 있다.

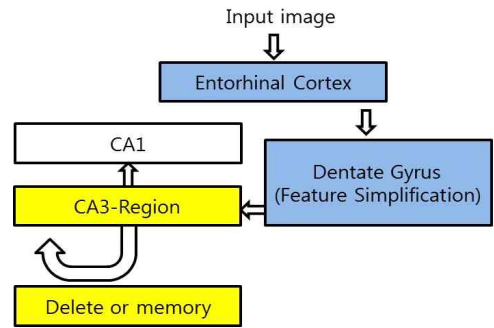


그림 4. 해마 시스템의 구조
Fig. 4. Structure of hippocampus system

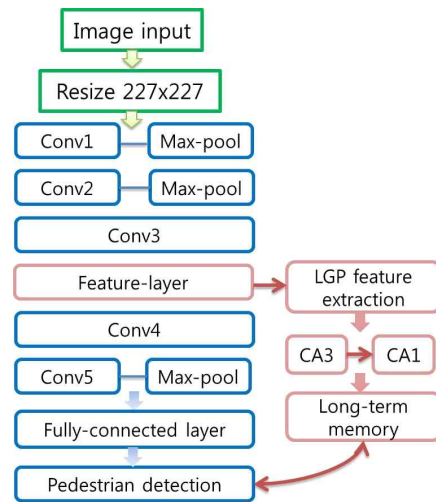


그림 5. 제안하는 알고리즘
Fig. 5. Proposed algorithm

III. 제안하는 방법

그림 5는 제안하는 CNN 구조를 나타낸다. 먼저 입력 영상의 크기를 227x227 영상으로 동일하게 변환한다. 그 후 총 5개의 층의 Convolution layer들을 거쳐 특징을 추출한다. 그 과정에서 Feature layer를 추가하였다. Feature layer에서는 LGP 특징 패턴을 연산하여 출현이 잦은 특징 패턴들을 장기 기억 장치에 저장한다. LGP 특징 패턴은 총 16가지의 패턴으로 변환되며 상위 빈도수 8가지 패턴을 저장시켜 검출에 사용 하였다.

Pooling은 Max-pooling을 적용했고, 첫 번째와 두 번째 그리고 마지막 Convolution layer에 적용시켰다. 그 후 Fully-Connected layer를 거쳐 학습 및 검출이 이루어진다.

장기 기억 패턴의 경우 학습 과정에서는 빈도 순서로 저장만을 수행하며 이후 Test Image를 이용한 검출 과정에서 저장된 패턴과 비교하는 과정을 통해 최종적으로 검출이 이루어진다.

3.1 Dataset 구성

Dataset은 표 1과 같이 INRIA Person dataset을 사용하였고 Train data는 총 2446장이 사용되었고 테스트 데이터는 총 741장이 사용되었다. Dataset에는 그림 6과 같이 Positive image와 Negative image가 포함되어 있다. 또한 그림 7과 같이 Train data를 좌우 반전 시킨 데이터를 추가하면서 Data augmentation 기법을 적용하였다.

표 1. 실험 dataset

Table 1. Experimental dataset

INRIA	Positive image	Negative image
Train	1228	1218
Test	288	453



그림 6. Positive image와 Negative image
Fig. 6. Positive image and negative image



그림 7. 좌우반전 시킨 Data
Fig. 7. Reverse data left and right

3.2 AlexNet 구조 최적화

본 논문에서 사용한 CNN 구조는 비교적 얇은 구조로 이루어진 AlexNet이다. 보행자 검출 문제의 경우 분류 기준이 보행자 한 가지이기 때문에 비교적 단순한 구조인 AlexNet 구조를 사용하여도 충분한 성능을 보인다. 기존의 AlexNet은 ImageNet에 있는 100 만장 이상의 이미지를 학습한 구조인데 반해 본 논문에서는 비교적 적은 학습 데이터에 맞게 AlexNet의 구조를 더 가볍게 변경하였다.

변경한 부분은 Convolution layer의 Feature map의 개수와 Fully-connected layer의 총 연결 Node 개수이다. 실험을 통해 최적의 구조를 찾는 과정을 거쳤다. CNN에서 대부분의 연산은 Convolution layer에서 이루어지기 때문에 검출률을 보존할 수 있다면 구조를 가볍게 할 필요가 있다. 전체 구조는 C1-C5 layer와 FC6-FC7 layer로 표현할 수 있다.

표 2. 변경된 CNN 구조

Table 2. Modified CNN architecture

C1	C2	C3	C4	C5	FC6	FC7	Detection rate(%)
96	256	384	384	256	4096	4096	95.4
48	128	192	192	256	2048	2048	95.3

표 2에서 기존의 구조와 변경된 구조를 나타내었다. 기존 AlexNet의 구조는 각 Convolution layer의 Feature map의 개수가 C1부터 C5까지 96, 256, 384, 384, 256개로 구성되어 있고 Fully connected layer의 node 개수는 FC6과 FC7 각각 4096개로 이루어져 있다. 구조를 가볍게 하기 위해 C5를 제외한 나머지 Layer들의 Feature map 개수와 node 개수를 1/2 만큼 줄여서 구성하였다[13]. 변경된 구조로 학습한 결과 표 2와 같이 검출률이 유지되는 것을 확인하였고 변경된 구조를 기준 모델로 사용한다. 그리고 기준 모델을 기반으로 Feature map과 연결 node 개수를 더 줄이면서 검출률의 변화를 실험하였다.

표 3. 변경된 CNN 구조와 검출률

Table 3. Modified CNN architecture and detection rate

C1	C2	C3	C4	C5	FC6	FC7	Detection rate(%)
48	128	192	192	256	2048	2048	95.3
24	64	96	96	128	1024	1024	93.5
24	128	192	192	256	2048	2048	94.5
48	64	192	192	256	2048	2048	94.7
48	128	96	96	256	2048	2048	93.1
48	128	192	192	128	1024	1024	92.8

표 3에서 기준 모델 보다 더 가벼운 구조의 수치와 검출률의 변화를 나타내었다. 우선 기준 모델에서 다시 1/2 만큼 개수를 줄여 실험한 결과 검출률이 약 2% 떨어졌다. 그 후 C1부터 C5까지 차례대로 개수를 줄여가며 실험하였고 실험 결과 검출률을 유지하는 더 가벼운 구조는 찾을 수 없었다. 더 많은 경우의 수가 있겠지만 모든 경우의 수를 실험해 보는 것은 과정이 너무 길어지기 때문에 처음 정한 기준 모델을 적용하였다.

3.3 S-LGP(Symmetry-LGP)

기본적인 LGP는 현재 픽셀과 이웃의 8개 픽셀을 비교하여 패턴이 생성되는데 이 값은 총 256가지로 나누어진다. 본 논문에서는 패턴을 더 단순화 시켜 8방향 모두 비교하지 않고 서로 마주 보는 픽셀들만 비교하여 총 16가지의 특징 패턴으로 간소화 시켰다[14]. S-LGP의 구조는 그림 8과 같이 나타낼 수 있다.

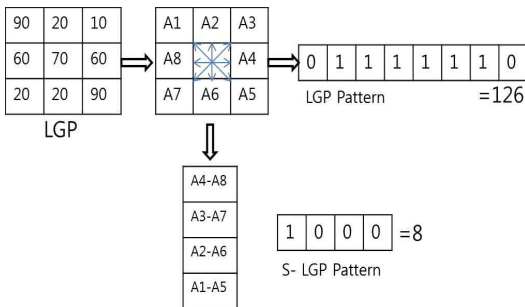


그림 8. S-LGP의 구조
Fig. 8. Structure of S-LGP

IV. 실험 결과

실험은 Geforce GTX980 Ti GPU를 이용해 Python 기반의 Tensorflow framework를 사용하여 진행하였다. 학습 알고리즘에는 딥 러닝 학습에서 자주 사용되는 stochastic gradient descent 방법[15]이 사용되었다. 실험을 통해 적절한 학습 epoch는 50으로 정하였다.

표 4. 데이터 전처리에 대한 비교
Table 4. Data preprocessing comparison

Preprocessing	Detection rate (%)
1-channel gray-scale	92.2
3-channel RGB	94.5
3-channel RGB + data augmentation	95.3

표 4는 3장에서 언급했던 Data augmentation에 대한 효과를 비교하여 나타낸 표이다. 먼저 입력 영상을 1채널 Gray-scale로 변환하여 실험한 결과 검출률이 약 3% 감소하였다. 그 후에 입력 영상을 3채널 RGB 채널 그대로 실험하였고 좌우 반전 데이터

를 약 600장 추가하여 학습한 결과 검출률이 더 개선되었다. 배경의 밝기 및 보행자의 각도 조절 등을 통해 더 많은 양의 Data augmentation을 추후 적용하면 더 개선된 검출률을 얻을 수 있을 것으로 보인다.

표 5. LGP-FL 적용 후 학습속도 및 검출률 비교
Table 5. Comparison of learning rate and detection rate after applying LGP-FL

Model	Training time (sec / batch)	Detection rate (%)
AlexNet	0.089	94.2
LGP-FL+AlexNet	0.127	95.3
S-LGP-FL+AlexNet	0.113	95.2

표 5에서는 제안하는 방법인 LGP-FL을 적용하였을 때의 학습 속도 및 검출률을 비교하였다. 기존의 AlexNet보다 검출률이 약 1% 더 개선되었음을 알 수 있다. 그러나 학습 속도에서는 중간에 LGP 특징 패턴을 연산하는 단계가 포함되어 약간의 손해가 있음을 알 수 있다. 추가로 장기 기억 패턴에 기존 LGP 특징의 256가지 패턴을 적용할 때보다 패턴을 간소화 시킨 S-LGP의 16가지 패턴을 적용시켰을 때 연산량을 줄이는 효과를 얻어 학습 속도를 줄일 수 있고 검출률에는 큰 변화가 없음을 알 수 있다.

그림 9는 제안하는 방법을 통한 보행자 검출의 결과이다. 단일 보행자부터 다중 보행자까지의 검출 결과를 나타내었다.

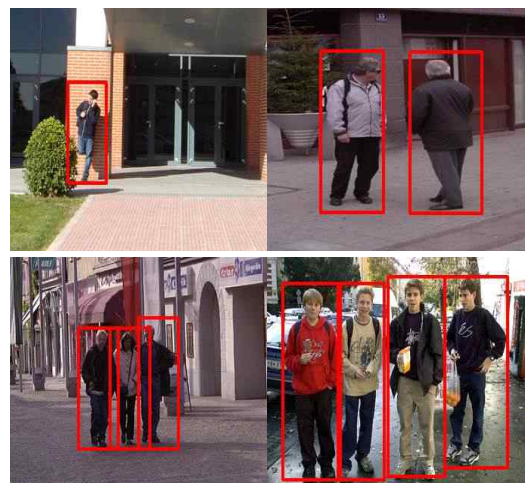


그림 9. 보행자 검출 결과
Fig. 9. Pedestrian detection results



(a)



(b)

그림 10. HOG 방법과 제안하는 기법의 검출 결과
Fig. 10. Detection results of HOG technique and proposed method

그림 10은 기존의 Hand-crafted 방법 중 보행자 검출에 주로 사용되는 특징 추출 방법인 HOG 특징 추출 방법의 검출 결과와 제안하는 기법의 검출 결과이다. 그림 10(a)에서 보듯이 HOG 특징을 이용한 보행자 검출에서는 배경의 색상과 보행자의 색상이 비슷할 경우에는 미검출이 일어나는 경우가 있었다. HOG 특징의 경우 주로 색상 차이에 의한 기술기를 기반으로 구성되어 있기 때문에 보행자와 배경 간의 색상 차이가 비슷한 경우에는 제대로 검출이 되지 않는 경우가 많다. 제안하는 기법을 통한 검출 결과 그림 10(b)와 같이 기존의 방법에서 미검출된 보행자까지 검출하는 것을 확인할 수 있다.

그림 11은 LGP-FL를 통한 해마 구조의 적용 유무에 따른 검출 결과를 나타낸다. 그림 11(a)의 경우 배경이 어두운 환경에서 기존의 AlexNet 모델을 사용하였을 때 검출되지 못한 경우였는데 그림 11(b)의 경우 LGP-FL를 적용시킨 모델에서 특징 패턴의 기억을 통해 보행자 내부의 특징패턴을 이용해 추가로 검출하는 것을 알 수 있었다. 어두운 환경에서도 검출이 가능한 이유는 학습을 할 때 Positive data에도 어두운 환경의 보행자가 존재하는데 그 데이터의 내부 특징 패턴을 이용해서 검출이 가능하게 된다. 표 6은 기존의 방법들과 제안하는 기법의 검출률 비교를 나타낸다.



(a)



(b)

그림 11. LGP-FL 적용한 검출 결과
Fig. 11. Detection results applied to LGP-FL

표 6. 검출 결과 비교

Table 6. Comparison of detection results

Methods	Detection rate(%)
HOG	90.48
HOG+Cascade	91.55
CNN (AlexNet)	94.18
Proposed Methods	95.25

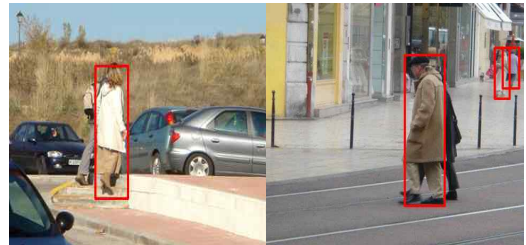


그림 12. 가려짐에 의한 미검출
Fig. 12. Undetected by occlusion

그림 12는 보행자 검출의 대표적 어려움 중 하나인 가려짐에 의한 미검출의 경우이다. 가려짐의 경우 해마 구조를 적용한 모델에서도 검출이 불가능하였다.

V. 결 론

본 논문에서는 대표적인 CNN 구조 중에 하나인 AlexNet을 Dataset의 크기에 맞게 재구성하고 그 구

조를 기반으로 LGP-FL을 추가하여 해마 구조를 적용한 보행자 검출 시스템을 구현하였다. 제안하는 기법을 이용하여 기존의 보행자 검출에서 좋은 성능을 가졌던 HOG+SVM과 Cascade HOG, 그리고 AlexNet 기반의 CNN보다 뛰어난 검출률을 얻을 수 있었다. 특히 학습 중간 단계에서 S-LGP 특징을 이용하여 보행자의 특징 패턴을 추출한 다음 빈도순서로 패턴을 저장하여 해마 구조의 장기 기억 장치를 구현하였다. 빈도순서로 저장되어 있던 보행자의 특징 패턴을 검출 시 이용하여 기존의 방법들에서 검출에 실패한 보행자들을 재검출하여 검출률을 약 1~4%까지 증가 시킬 수 있었다. 기존의 방법들에서는 보행자와 배경과의 색상 차이, 어두운 환경의 영상 등에서 검출 실패가 있었지만 밝기 변화에 강인한 LGP 특징 패턴을 이용하여 추가 검출이 가능하게 하였다. 그러나 학습 및 검출 속도는 특징 패턴의 연산이 추가됨으로써 약간 증가하였음을 알 수 있다. 추후 실험을 통해 검출률은 유지하면서 속도를 더 빠르게 하기 위해 더 최적화된 CNN 구조를 구성해볼 필요가 있다. 추후 연구에서는 가려진 보행자를 검출할 방안을 찾고 실시간 영상에서 검출 후 추적까지 가능한 방향으로 연구를 진행할 계획이다.

References

- [1] P. Viola and M. J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p. 9, Dec. 2001.
- [2] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions", Pattern Recognition, Vol. 29, No. 1, pp. 51-59, Jan. 1996.
- [3] Dalal and B. Triggs, "Histograms of oriented gradients for human detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 886-893, Jun. 2005.
- [4] Q. Zhu, M. C. Yeh, K. T. Cheng, and S. Avidan, "Fast human detection using a cascade of histograms of oriented gradients", In Proc. CVPR, Vol. 2, pp. 1491-1498, Jun. 2006.
- [5] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks", Advances in neural information processing systems, Vol. 1, pp. 1097-1105, Dec. 2012.
- [6] Simonyan, Karen and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition", Published as a conference paper at ICLR 2015, arXiv preprint arXiv:1409.1556, Sep. 2014.
- [7] Szegedy, Christian, and et al, "Going deeper with convolutions", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Sep. 2014.
- [8] Seok-Gyu Choi and Wenjie Xu, "A Study on Person Re-Identification System using Enhanced RNN", JIIBC, Vol. 17, No. 2, pp. 15-23, Apr. 2017.
- [9] Bongjin Jun and Daijin Kim, "Robust face detection using local gradient patterns and evidence accumulation", Pattern Recognition, Vol. 45, No. 9, pp. 3304-3316, Sep. 2012.
- [10] Kiho Kong and Dae-Seong Kang, "A Study of Face Detection Algorithm Using CNN with Mixed-LGP and Hippocampus Structure", Journal of KIIT, Vol. 16, No. 1, pp. 11-17, Jan. 2018.
- [11] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition", Proceedings of the IEEE, Vol. 86, No. 11, pp. 2278-2324, Nov. 1998.
- [12] P. Dayan and L. F. Abbott, "Theoretical Neuroscience", MIT press, 2001.
- [13] In-kyu Choi, Hyok Song, Sangyong Lee, and Jisang Yoo, "Facial Expression Classification Using Deep Convolutional Neural Network", JBE Vol. 22, No. 2, pp. 162-172, Mar. 2017.
- [14] Ki-Ho Kong and Dae-Seong Kang, "A Study of Face Detection Algorithm Using CNN Based on

Symmetry-LGP & Uniform-LGP and the Skin Color", Journal of KIIT. Vol. 15, No. 1, pp. 107-113, Jan. 2017.

- [15] Vogl, P. Thomas, et al., "Accelerating the convergence of the back-propagation method", Biological cybernetics, Vol. 59, No. 4, pp. 257-263, Sep, 1988.

저자소개

박 수 빈 (Su-Bin Park)



2015년 8월 : 동아대학교
전자공학과 (공학사)
2015년 8월 ~ 현재 : 동아대학교
전자공학과 석사과정
관심분야 : 컴퓨터비전, AI

강 대 성 (Dae-Seong Kang)



1994년 5월 : Texas A&M 대학교
전자공학과(공학박사)
1995년 ~ 현재 : 동아대학교
전자공학과 교수
관심분야 : 영상처리, 패턴인식