

특집논문 (Special Paper)

방송공학회논문지 제23권 제3호, 2018년 5월 (JBE Vol. 23, No. 3, May 2018)

<https://doi.org/10.5909/JBE.2018.23.3.345>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

CNN을 이용한 거리 사진의 분류와 안전도 평가

배 규 호^{a)}, 윤 정 언^{a)}, 박 인 규^{a)†}

Classification and Safety Score Evaluation of Street Images Using CNN

Kyu Ho Bae^{a)}, Jung Un Yun^{a)}, and In Kyu Park^{a)†}

요 약

CNN (convolutional neural network)은 최근 가장 주목받는 인공지능 기법 중 하나이며 특히 영상 분류에서 기존의 기법에 비해 월등한 성능을 보인다. 본 논문에서는 CNN을 이용하여 다양한 거리 사진을 분류하고, 분류 결과를 이용하여 해당 거리에 대한 안전도의 평가 방법을 제안한다. 제안하는 기법은 CNN을 이용하여 총 네 가지 유형의 거리 사진에 대하여 학습을 수행하는 과정과 학습된 네트워크 모델을 바탕으로 해당 거리 사진의 분류와 안전도를 평가하는 과정을 포함한다. 거리 사진의 학습 과정에서는 네 가지 유형의 거리 사진 데이터셋을 수집하고 이 데이터를 증강시킨 후 CNN 학습을 수행한다. 학습된 CNN 모델은 주어진 입력 영상의 분류를 정확히 수행하고, 거리의 안전도는 각 유형에 대한 확률을 조합하여 정량적으로 계산한다.

Abstract

CNN (convolution neural network) has become the most popular artificial intelligence technique and shows remarkable performance in image classification task. In this paper, we propose a CNN-based classification method for various street images as well as a method of evaluating the safety score for the street. The proposed method consists of learning four types of street images using CNN and classifying input street images using the learned CNN model followed by evaluating the safety score. During the learning process, four types of street images are collected and augmented, and then CNN learning is performed. It is shown that learned CNN model classifies input images correctly and the safety scores are evaluated quantitatively by combining the probabilities of different street types.

Keywords : CNN, street photo, classification, safety score

a) 인하대학교 정보통신공학과(Inha University, Department of Information and Communication Engineering)

† Corresponding Author : 박인규(In Kyu Park)

E-mail: pik@inha.ac.kr

Tel: +82-032-860-9190

ORCID:<https://orcid.org/0000-0003-4774-7841>

※ 이 논문은 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2016R1A2B4014731).

· Manuscript received March 16, 2018; Revised May 8, 2018; Accepted May 8, 2018.

1. 서 론

거리의 외관과 범죄 발생률 사이의 연관성에 대한 연구로 널리 알려진 ‘깨진 유리창 이론’^[1]에 따르면 깨진 유리창 하나로 범죄가 확산되어 사회적 문제로 확대될 수 있다. 이와 같이 거리의 외관과 범죄 발생률 사이에는 큰 관련이 있다. 대다수의 사람은 거리의 사진을 보고 그곳이 안전한

지 그렇지 않을지를 직관적으로 판단할 수 있지만 컴퓨터가 이를 정확히 판단하는 것은 쉽지 않은 문제이다. 이 문제의 해결을 통하여 영상 기반 보안, 디지털 지도 서비스, 청소년 계도 등 다양한 응용 분야에서 응용될 수 있을 것이다.

Streetscore는 γ -support vector regression (γ -SVR)과 Microsoft Trueskill 알고리즘을 활용하여 다양한 장소의 사진에 대해 안전도를 결정짓는 특징들을 추출한 뒤 Q-점수의 형태로 안전도 점수를 계산하고 이를 도시 규모의 지도에 시각화 하였다^[2]. Ordonez 등의 연구에서는 이를 안전도 예측 뿐 아니라 지역적 특이성 및 부유한 정도의 예측까지 확장하였다^[3]. Arietta 등은 범죄율과 연관이 있는 거리의 시각적 요소를 추출하여 해당 거리의 범죄율을 예측하는 연구를 수행하였고^[4], Naik 등은 [2]를 기반으로 거리 외관의 변화를 감지하는 알고리즘을 제안하였다^[5]. Traunmueller 등은 거리의 외관이 아닌 거리를 지나며 마주치는 사람에 따라 해당 거리의 안전도를 평가하는 알고리즘을 제안하였다^[6].

본 논문에서 제안하는 기법은 이러한 기존 연구를 기반으로 최근 영상 분류 문제에 뛰어난 성능을 보이고 있는 심층 학습모델의 하나인 합성곱 신경망(convolutional neural network, CNN)을 이용한다. 본 논문에서는 CNN을 이용하여 다양한 거리 사진에 대한 학습을 수행한 뒤 이를

활용하여 임의의 거리에 대한 안전도 점수를 평가한다.

II. 제안하는 기법

본 논문에서 제안하는 기법은 기존의 CNN 기반 학습 기법을 이용하여 네 가지 분류 유형의 거리에 대한 학습을 수행하는 단계와 학습된 네트워크를 이용하여 입력 사진을 분류하고 보행자와 운전자 기준의 거리 안전도를 평가하는 두 단계로 구성된다.

1. CNN 학습 방법

본 논문에서는 학습할 데이터 셋의 크기가 작고 분류할 범주의 수가 적기 때문에 CNN 초기 모델인 AlexNet [7]을 변형 없이 사용하여도 충분한 분류 성능을 얻을 것으로 가정한다. 학습할 네 가지 거리 유형은 도시의 거리에서 흔히 볼 수 있는 ‘공사현장’, ‘비포장 도로’, ‘사람이 많이 다니는 길’, ‘넓은 차도’이다. 각각의 학습 데이터는 AlexNet의 기본 영상 해상도인 227×227 크기로 바꾸어 전체 4,868장의 데이터를 7:3의 비율로 나누어 학습과 확인(validation) 데



그림 1. 학습에 사용된 데이터 예시. (a) 공사현장, (b) 비포장 도로, (c) 사람이 많이 다니는 길, (d) 넓은 차도

Fig. 1. Examples of images used for training. (a) construction site, (b) unpaved road, (c) street with many people, (d) broad road

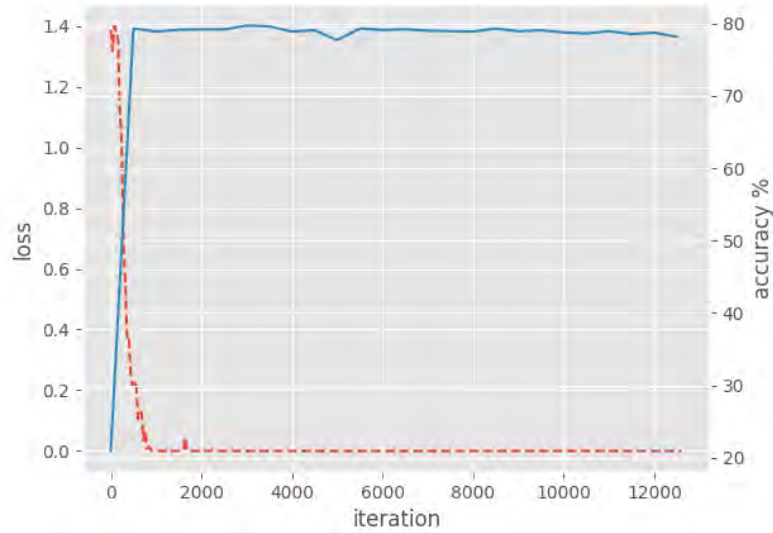


그림 2. 제안하는 기법의 CNN 학습 결과 (붉은 점선: loss, 파란 실선: accuracy)
Fig. 2. CNN training result of the proposed method (Red dashed: loss, Blue solid: accuracy)

이터로 활용하여 학습을 수행한다. 학습에 사용한 학습률(learning rate)은 0.005로 설정하였고 256의 배치 크기(batch size)로 학습을 약 12600회 반복하였다. 확인 과정은 학습을 500번 반복할 때 마다 50의 배치 크기로 500회씩 반복하여 수행하였다.

학습에 사용된 거리 사진은 Daum 카카오사의 로드뷰 거리영상을 캡처하여 취득하였으며, 그림 1에 그 사례를 제시하였다.

학습 데이터에 대해 CNN 학습을 수행한 결과는 그림 2와 같다. 붉은 점선은 학습과정에서의 손실(loss)값이며, 파란 실선은 확인 데이터셋으로 측정한 정확도를 나타낸다. 그래프를 살펴보면 손실 값이 줄어들고 동시에 정확도가 상승하는 것을 확인할 수 있고, 두 값 모두 적정선에서 수렴하며 약 80%의 정확도를 갖는 것을 확인할 수 있다.

2. 보행자 기준 안전도 평가

앞 단계에서 학습된 CNN을 이용하여 임의의 입력 거리 사진을 네 가지 유형중 하나로 분류한 뒤 다음 식과 같이 보행자 기준 안전도 점수(S_p)를 부여 한다.

$$S_p = \left[\frac{2P_{pedestrian} + 0.5P_{broad} - P_{construction} - P_{unpaved}^{*1}}{3} \right] \times 5 \quad (1)$$

$$S_d = \left[\frac{2P_{broad} - P_{pedestrian} - 0.7P_{unpaved} - 0.5P_{construction} + 1}{3} \right] \times 5 \quad (2)$$

식 (1)의 P는 각 거리 유형별 확률을 의미한다. CNN이 결과 값으로 출력하는 Softmax 확률 값을 사용하여 안전도 점수를 최소 0점부터 최대 5점까지의 범위로 수치화 한다. 각 거리 유형별 확률에 부여한 가중치는 경험적으로 결정하였다. 즉, 보행자의 관점에서 ‘사람이 많이 다니는 길’의 경우 안전하다고 느끼기 때문에 2의 가중치를 부여한다. ‘넓은 차도’는 조명이 양호하고 통행하는 차량의 수가 많기 때문에 비교적 안전하다고 할 수 있다. 그러나 도로교통공단의 교통사고 분석시스템^[8]에 의하면 차대사람 사고 중 도로 횡단과 차도 통행 관련 사고가 63%정도를 차지하기 때문에 0.5의 가중치를 부여한다. 나머지 두 가지 유형의 경우 기존의 연구^[2]와 같이 음의 점수를 부여한다.

3. 운전자 기준 안전도 평가

보행자의 시각으로 본 거리의 안전도와 운전자의 시각으로 본 거리의 안전도 사이에는 차이가 있기 때문에 다음과 같이 보행자 기준 안전도 점수와 다른 안전도 점수를 부여한다.

도로교통공단의 교통사고 분석시스템^[8]에 따르면, 전체 교통사고 합계 중 13m이상의 차도 폭을 갖는 경우 사고 발

생물이 약 29%정도이고 9m미만의 경우 약 54%의 비중을 차지한다. 따라서 ‘넓은 차도’에 2.0의 가중치를 부여하고 ‘비포장 도로’의 경우 대부분 폭이 좁기 때문에 -0.7의 가중치를 부여한다. ‘사람이 많이 다니는 길’의 경우 운전자의 관점에서 사람이 많은 경우 차대사람 교통사고의 위험이 크므로 -1의 가중치를 부여하고, ‘공사현장’의 경우 해당 유형의 특징상 돌발 상황이 자주 발생하므로 -0.5의 가중치를 부여한다.

III. 실험 결과

1. 분류 유형 영상에 대한 실험 결과

본 논문의 실험은 Intel i5-6500 3.6GHz CPU와 NVIDIA GeForce GTX 1060 6GB를 장착한 컴퓨터에서 수행하였고 실험에 이용한 영상은 Daum 카카오사의 로드뷰에서 취득하였다.



그림 3. Test 데이터셋
Fig. 3. Test dataset

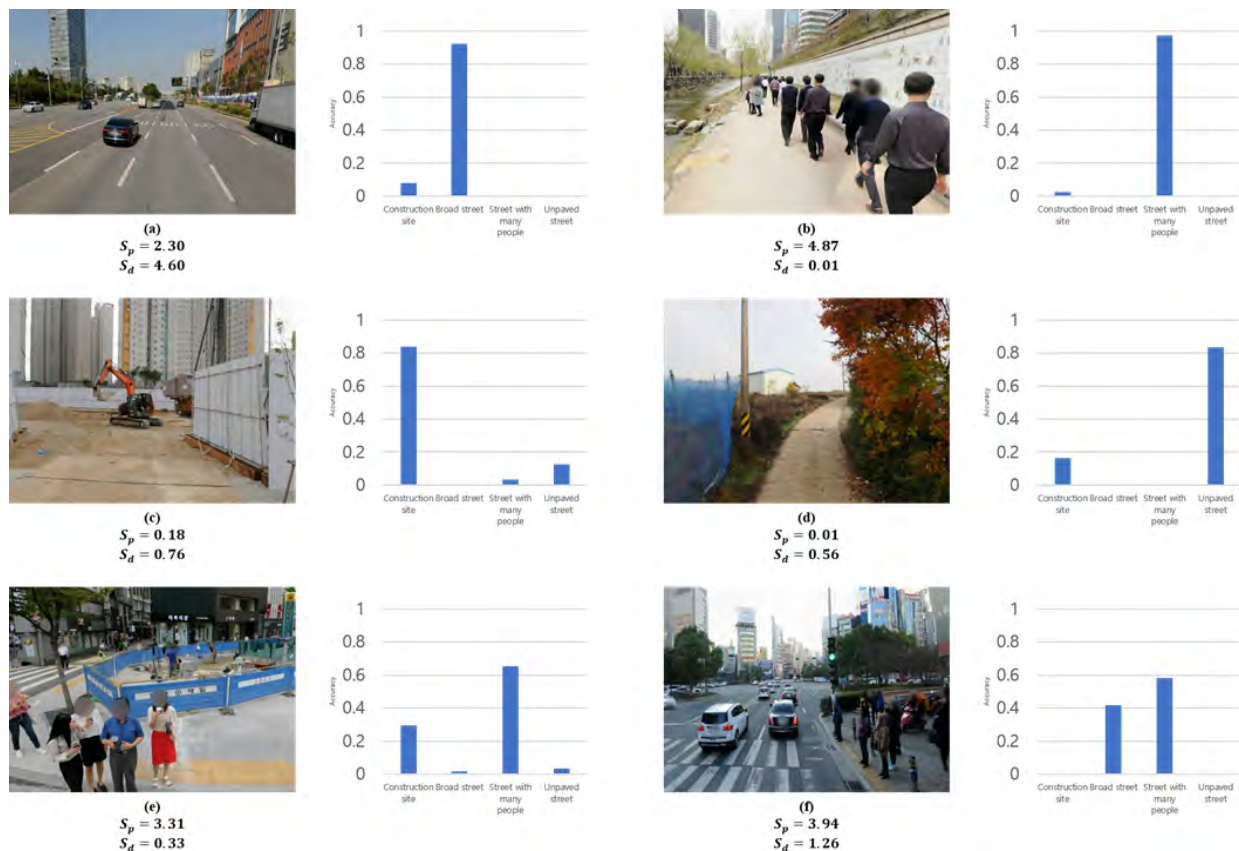


그림 4. 실험영상에 대한 분류결과 및 안전도 점수 예시
Fig. 4. Classification results and safety scores for test examples

실험에 사용한 학습 결과를 확인하기 위해 각 유형별로 20장 내외의 사진을 수집하고 추가적으로 두 가지 이상의 유형이 혼합되어 있는 사진을 10여장 수집하여 그림 3과 같이 총 113장의 사진에 대하여 실험을 수행하였다. 데이터 셋의 6개의 예시 사진에 대한 분류 결과와 안전도 점수를 구하여 그림 4에 제시하였다. 그림 4(c)와 그림 4(d)와 같은 공사현장이나 비포장 도로의 경우 보행자와 운전자 기준 모두 낮은 안전도 점수를 보인다. 그림 4(a)와 같이 넓은 차도의 경우 보행자 기준 안전도 점수가 상대적으로 낮음을 확인할 수 있고, 그림 4(b)에서는 사람이 많은 길의 경우 운전자 기준 안전도 점수가 상대적으로 낮음을 확인할 수 있다. 그림 4에서 볼 수 있듯이 사람이 분류할 때 얻을 수 있는 결과와 동일한 분류 범주에서 가장 높은 확률 값을 보인다.

전체 실험 영상의 ground truth 대비 분류 성공률을 표 2에 제시하였다. 본 실험에서는 네 가지 유형의 실험 영상의 경우 분류 결과가 70% 이상인 경우를 성공으로 정의한다. 혼합 유형의 경우 분류 결과가 해당 영상에 포함 된 두 가지 유형을 모두 나타내고 영상에서의 특징 비율에 따라 한 가지 유형이 최대 80%를 넘지 않는 경우를 성공으로 정의한다. 표 2에서 볼 수 있듯이 전체 분류 범주 중 ‘비포장 도로’의 분류 성공률이 다소 낮은 반면에, ‘공사현장’의 분류 성공률이 다소 높은 것을 확인할 수 있다. ‘비포장 도로’의 경우 ‘공사현장’과 유사한 특징을 갖고 있고 따라서 이를 ‘공사현장’으로 잘못 분류하는 현상이 종종 발생하여 이러한 결과가 나타남을 확인하였다. 전체 실험 영상에 대해서는 84.1%의 성공률을 보였다.

표 2. 전체 실험 영상에 대한 분류 성공률
Table 2. Correct classification rate for the entire experimental dataset

	Construction Site	Broad Street	Street with Many People	Unpaved Street	Mixed Category	Average
Accuracy (%)	92.6	81.8	87.1	77.3	72.7	84.1

IV. 결 론

본 논문에서는 CNN 기반의 학습을 이용한 거리 유형의 분류 기법과 각 거리 사진에 대한 보행자, 운전자 기준 안전도 점수 평가 기법을 제안하였다. 제안한 방법은 네 가지 유형의 거리 사진에 대하여 CNN 기반의 학습을 통해 단일 유형의 거리 사진과 두가지 유형이 혼합된 사진에 대해서도 정확히 분류를 수행한다. 분류 결과를 바탕으로 각 거리 사진에 대해 보행자, 운전자 기준 안전도 점수를 평가할 수 있다. 향후 더 많은 학습 데이터와 보다 세분화된 거리 분류에 대해 본 연구를 확장하여 실용적인 시스템으로 확장할 수 있다.

참 고 문 헌 (References)

- [1] J. Q. Wilson and G. L. Kelling, "Broken windows," *Atlantic Monthly*, vol. 249, no. 3, pp. 29-38, March 1982
- [2] N. Naik, J. Philipoom, R. Raskar, and C. A. Hidalgo, "Streetscore-Predicting the perceived safety of one million streetscapes," *Proc. of IEEE CVPR Workshop*, pp. 779-785, 2014.
- [3] V. Ordonez and T. L. Berg, "Learning high-level judgments of urban perception," *Proc. of ECCV*, pp. 494-510, 2014
- [4] S. M. Arietta, A. A. Efros, R. Ramamoorthi, and M. Agrawala, "City forensics: Using visual elements to predict non-visual city attributes," *IEEE Trans. on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2624-2633, December 2014
- [5] N. Naik, S. D. Kominers, R. Raskar, E. L. Glaeser, and C. A. Hidalgo, "Computer vision uncovers predictors of physical urban change," *National Academy of Sciences*, vol. 114, no. 29, pp. 7571-7576, July 2017
- [6] M. Traunmueller, P. Marshall, and L. Capra, "...when you're a Stranger: Evaluating Safety Perceptions of (un) familiar Urban Places," *Proc. of Second International Conference on IoT in Urban Space*, pp. 71-77, 2016
- [7] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolution neural networks," *Proc. of NIPS*, pp. 1097-1105, 2012.
- [8] KoROAD, *Traffic Accident Analysis System*, 2016

저 자 소 개



배 규 호

- 2018년 2월 : 인하대학교 정보통신공학과 학사
- 2018년 3월 ~ 현재 : 인하대학교 정보통신공학과 석사과정
- ORCID : <https://orcid.org/0000-0003-1239-8522>
- 주관심분야 : 컴퓨터비전, deep learning, GPGPU



윤 정 언

- 2018년 2월 : 인하대학교 정보통신공학과 학사
- 2018년 3월 ~ 현재 : 인하대학교 정보통신공학과 석사과정
- ORCID : <https://orcid.org/0000-0001-6091-366X>
- 주관심분야 : 컴퓨터비전, deep learning, GPGPU



박 인 규

- 1995년 2월 : 서울대학교 제어계측공학과 학사
- 1997년 2월 : 서울대학교 제어계측공학과 석사
- 2001년 8월 : 서울대학교 전기컴퓨터공학부 박사
- 2001년 9월 ~ 2004년 3월 : 삼성종합기술원 멀티미디어랩 전문연구원
- 2007년 1월 ~ 2008년 2월 : Mitsubishi Electric Research Laboratories (MERL) 방문연구원
- 2014년 9월 ~ 2015년 8월 : MIT Media Lab 방문부교수
- 2004년 3월 ~ 현재 : 인하대학교 정보통신공학과 교수
- ORCID : <http://orcid.org/0000-0003-4774-7841>
- 주관심분야 : 컴퓨터 그래픽스 및 비전 (영상기반 3차원 형상 복원, 증강현실, computational photography), GPGPU