Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography
Vol. 38, No. 3, 269-279, 2020
https://doi.org/10.7848/ksgpc.2020.38.3.269

ISSN 1598-4850(Print)
ISSN 2288-260X(Online)
Original article

스테레오 CCTV 영상에서 딥러닝을 이용한 교통량 추정

Estimation of Traffic Volume Using Deep Learning in Stereo CCTV Image

서홍덕[1] · 김의명[2]
Seo, Hong Deok · Kim, Eui Myoung

Abstract

Traffic estimation mainly involves surveying equipment such as automatic vehicle classification, vehicle detection system, toll collection system, and personnel surveys through CCTV (Closed Circuit TeleVision), but this requires a lot of manpower and cost. In this study, we proposed a method of estimating traffic volume using deep learning and stereo CCTV to overcome the limitation of not detecting the entire vehicle in case of single CCTV. COCO (Common Objects in Context) dataset was used to train deep learning models to detect vehicles, and each vehicle was detected in left and right CCTV images in real time. Then, the vehicle that could not be detected from each image was additionally detected by using affine transformation to improve the accuracy of traffic volume. Experiments were conducted separately for the normal road environment and the case of weather conditions with fog. In the normal road environment, vehicle detection improved by 6.75% and 5.92% in left and right images, respectively, than in a single CCTV image. In addition, in the foggy road environment, vehicle detection was improved by 10.79% and 12.88% in the left and right images, respectively.

Keywords : Artificial Intelligence, Deep Learning, CCTV, Traffic Volume, Affine Transformation, Fog

초 록

교통량 산정은 주로 교통량조사시스템, 차량검지시스템, 통행료징수시스템 등과 같은 조사 장비와 CCTV를 통한 인력 조사를 병행하고 있으나 이는 많은 인력과 비용이 발생한다. 본 연구에서는 단일 CCTV의 경우 전체 차량을 탐지하지 못하는 한계를 극복하기 위해서, 딥러닝과 스테레오 CCTV를 이용하여 교통량을 산정하는 방법을 제안하였다. 차량을 탐지하기 위한 딥러닝 모델을 학습하기 위해 COCO 데이터셋을 사용하고, 실시간으로 좌우 CCTV 영상에서 각각 차량을 탐지하였다. 그리고 나서, 각 영상에서 추출하지 못한 차량을 부등각사상변환을 이용하여 추가적으로 차량을 탐지하여 교통량 산정의 정확도를 개선하였다. 실험은 평상시 도로 환경과 안개가 발생한 기상 상황의 경우에 대해서 각각 수행하였다. 평상시 도로 환경의 경우 단일 CCTV 영상을 사용할 때보다 좌우 영상에서 각각 6.75%, 5.92%의 차량 탐지의 개선효과가 있었다. 또한, 안개가 발생한 도로 환경의 경우 좌우 영상에서 각각 10.79%, 12.88%의 차량 탐지의 개선효과가 있었다.

핵심어 : 인공지능, 딥러닝, CCTV, 교통량, 부등각사상변환, 안개

Received 2019. 12. 30, Revised 2020. 02. 10, Accepted 2020. 03. 09
1) Member, Department of Spatial Information Engineering, M.S. Student, Namseoul University (E-mail: seoing23@gmail.com)
2) Corresponding Author, Member, Department of Spatial Information Engineering, Professor, Namseoul University (E-mail: kemyoung@nsu.ac.kr)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서 론

최근 컴퓨터의 그래픽 가속을 담당하는 장치인 GPU (Graphics Processing Unit)와 같은 하드웨어 성능의 향상으로 인해 카메라를 이용하여 수집하는 영상에서 딥러닝(deep learning)을 통해 객체를 탐지(detection)하고 객체의 영역을 분할(segmentation)하는 다양한 연구가 수행되고 있다(Jeong and Jeong, 2019; Lee and Yom, 2018; Young *et al*., 2018; Xu *et al*., 2018; Lee *et al*., 2018). 또한, 딥러닝과 도로가 촬영된 영상을 이용하여 차량을 탐지하는 연구 및 교통량 산정 연구도 활발히 진행 중이다(Ammour *et al*., 2017; Mundhenk *et al*., 2016; Du, *et al*., 2017).

국토교통부의 교통량정보제공시스템에 의하면 교통량 산정은 기본 교통량 자료가 필요하다고 판단되는 모든 구간에 대하여 광범위하게 실시하는 수시조사와 특정 지점에 자동차종 분류 조사 장비를 설치하여 1년 365일 24시간 연속으로 통과 차량의 차종, 방향, 시간대별로 측정하는 상시조사를 수행한다(Traffic Monitoring System, 2018).

수시조사와 상시조사는 교통량조사시스템(AVC: Automatic Vehicle Classification), 차량검지시스템(VDS: Vehicle Detection System), 통행료징수시스템(TCS: Toll Collection System) 등의 조사 장비와 CCTV (Closed Circuit TeleVision)를 통한 인력식 조사와 조사원을 조사지점에 배치하여 차종별, 방향별, 시간대별 교통량 조사를 병행하고 있다(Traffic Monitoring System, 2018).

CCTV는 교통에 대한 정성적인 정보를 판독할 수 있기 때문에 보급이 확산되고 있다(Jo *et al*., 2018). 그러나, CCTV의 수는 기하급수적으로 증가하면서 관제요원의 눈에만 의존하는 단순 모니터링 방식의 문제점들이 나타나기 시작하였으며, CCTV 관제를 위한 통합관제 센터는 1인당 평균 46대의 CCTV를 관리함으로서 중요 이벤트를 수작업으로 감지하기는 쉽지 않은 실정이다(Kim *et al*., 2018; Han *et al*., 2019; Sirirattanapol *et al*., 2019). 따라서, 영상분석을 통해 전반적인 상황을 인지하는 지능형 CCTV 기술이 필수적으로 필요하며 이를 위해 딥러닝을 이용한 차량 검출 알고리즘에 대한 연구가 진행되고 있다(Choi and Yoo, 2017; Lee *et al*., 2018).

CCTV를 이용한 교통량 산정의 선행 연구들을 살펴보면 차량에 2대의 CCTV를 장착하여 주행하는 방향과 반대 방향으로 주행하는 차량을 촬영하여 교통량을 산정하는 연구(Huh *et al*., 2013), CCTV 영상과 유사한 과거 문제의 솔루션을 기반으로 새로운 문제를 해결하는 CBM (Cased Based Management) 알고리즘을 이용하여 교통 혼잡을 분해하기 위한 솔루션 모델을 제안하는 연구(Anindra *et al*., 2018), 도로에 설치된 CCTV로부터 실시간으로 영상을 받아 입력된 영상에서 다양한 색상, 기하학적 특징 등을 추출하여 차량을 구별하고 이를 통해 차량의 통행량을 계산하는 연구(Hong *et al*., 2011) 등이 수행되었다.

CCTV 영상과 딥러닝을 이용하여 차량 탐지 및 교통량 산정 선행 연구들을 살펴보면 다양한 조명 조건의 영상에서 YOLO (You Only Look Once) 모델을 통해 차량과 사람을 탐지하는 연구(Tung *et al*., 2019), CCTV 영상의 차량을 SSD (Single Shot Multibox Detector) MobileNet V1과 Faster R-CNN (Convolutional Neural Network) ResNet101과 같은 사전 훈련된 모델을 통해 교통의 흐름을 추정하는 연구(Peppa *et al*., 2018), 딥러닝 객체 탐지 알고리즘인 Faster R-CNN을 이용하여 차량 객체를 탐지하고 추적하여 차종, 차량 통행량, 통행 속도를 추정하는 연구(Kim and Choi, 2019), CCTV 영상과 YOLO V3 모델을 통해 자전거, 차량, 보행자 등의 교통 흐름을 모니터링하기 위한 엣지 검출에 관한 연구(Barthélemy *et al*., 2019), CCTV 영상과 CNN을 이용해 기상 데이터, 교통 현황, 도로 표면 상태 등 운전자에게 필수적인 도로 환경 정보를 추정하는 연구(Sirirattanapol *et al*., 2019), CCTV 영상과 SSD를 이용해 도로환경 CCTV 영상에서 유효 학습데이터를 추출하고 이를 이용하여 실시간으로 동작하는 고성능의 객체검출 모델을 생성하는 연구(Choi and Yoo, 2017) 등이 수행되었다.

그러나, 딥러닝과 단일 CCTV 영상을 이용하여 차량을 검출할 경우 촬영 환경에 따른 노이즈 발생과 낮은 공간해상도로 인한 인식률의 저하가 발생한다. 이에 본 연구에서는 스테레오 CCTV 영상에 딥러닝 모델 중 하나인 YOLO V3를 활용하여 도로 및 구간별 교통량을 산정하는 방법을 제안하는 것을 목적으로 하였다.

2. 연구방법

2.1 제안한 방법론

본 연구에서는 도로의 교통량을 산정하기 위하여 스테레오 CCTV 영상에서 딥러닝을 이용하여 차량을 탐지하는 방법을 Fig. 1과 같이 제안하였다. COCO(Common Objects in Context) 데이터셋과 테스트할 스테레오 CCTV 영상을 각 프레임별로 구축한 후, COCO 데이터셋을 YOLO V3를 이용하여 학습시킨다.

그 다음으로 COCO 데이터셋을 학습한 YOLO V3를 이용하여 CCTV의 좌우 영상에서 차량을 각각 탐지한다. 탐지된

좌우 영상은 영상의 특징이 다를 수 있기 때문에 진단이나 예측의 정확도에 차이가 날 수 있다(Park and Bae, 2019). 따라서, 좌우 영상에서 탐지한 차량의 개수가 완전히 동일하지 않을 수 있기 때문에 좌우 영상 간의 부등각사상변환 매개변수를 추정한 후, 좌우 영상에서 탐지된 차량의 좌표와 추정한 부등각사상변환 매개변수를 이용하여 각 영상에서 탐지하지 못한 차량을 탐지하여 좌우 영상의 차량 대수를 일치시킨 후 교통량을 계산한다.

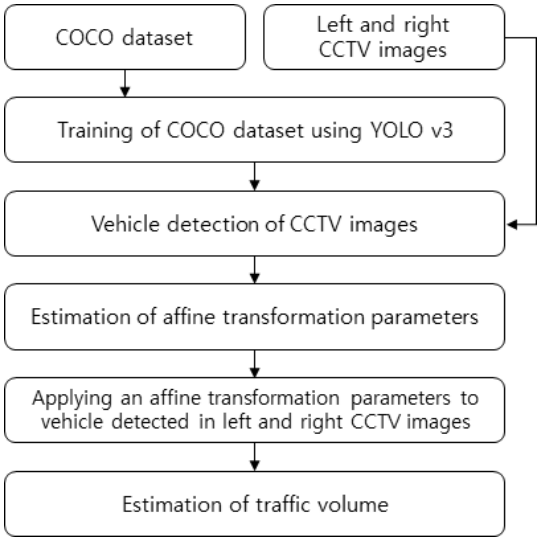


Fig. 1. Research flow chart

2.2 COCO Dataset

COCO 데이터셋은 마이크로소프트에서 객체 인식의 첨단화를 목표로 만든 딥러닝을 위한 데이터셋이다(Lin *et al*., 2014).

영상에서 딥러닝을 통해 차량을 탐지하기 위해서는 다양한 환경의 영상을 확보하고 영상에서 탐지하고자 하는 모든 객체를 Fig. 2와 같이 바운딩 박스(bounding box)로 표시해야 한다. 또한, 바운딩 박스 위치의 각 좌표를 기록해야하며, 지정된 클래스로 세분화하는 것이 필수적이다. 그러나, 이와 같이 차량 탐지를 위한 레이블링을 위해서 많은 시간, 비용, 인력 등이 필요하므로, 이러한 문제를 해결하기 위해서 기 구축되어 있는 COCO 데이터셋을 사용한다.

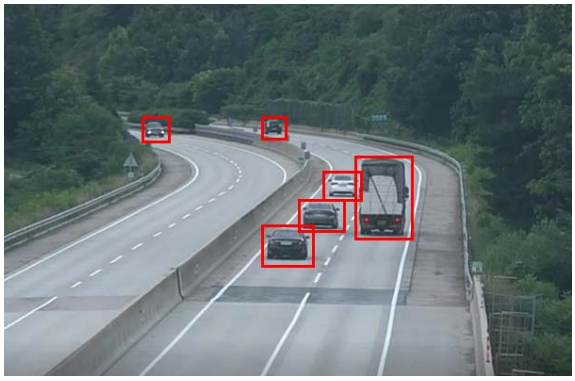


Fig. 2. Localization of the bounding boxes of objects

COCO 데이터셋은 객체 탐지, 객체 분할, 특징점 탐지를 위해서 약 30만 장 이상의 영상을 취득하여 객체를 91개의 클래스로 나누어 Table 1과 같이 약 25만 개의 레이블로 표시된 객체를 가지고 있으며, 본 연구에서는 차량을 탐지하기 위해 기 구축되어 있는 COCO 데이터셋의 차량 데이터를 활용하였다(Lin *et al*., 2014).

Table 1. COCO Dataset statistics

Category	Number of objects
ALL	245,496
TRAIN	82,783
TEST2015	81,434
TEST	40,775
VALIDATION	40,504

2.3 YOLO V3

YOLO V3 모델은 실시간 객체 탐지에 최적화된 모델로 다른 딥러닝 네트워크 모델에 비해 상당히 빠른 처리 속도를 보여준다(Lin *et al*., 2014). YOLO V3는 기존 YOLO 모델에서 소형의 물체가 겹쳐있는 경우 그 위치를 정확하게 추정하지 못하는 단점 등을 개선한 모델로 이를 보완하기 위해 Backbone Network는 DarkNet-53을 이용하였으며, 입력 영상을 다양한 축척으로 변환하여 각 축척에 탐지 커널을 이용해 객체를 탐지한다(Yu *et al*., 2019; Redmon and Farhadi, 2018).

YOLO V3의 학습은 Fig. 3과 같이 입력 영상을 $S \times S$ 크기의 격자로 나눈 후, 각 격자에 바운딩 박스와 스코어($score$) 값을 계산한다. 여기서, 빨간 격자는 분류할 차량이 존재하는 경우를 나타내며, 파란 격자는 차량이 존재하지 않은 경우를 나타낸다. 따라서, 스코어 값은 분류할 객체가 있을 확률이 얼마나 되는지 그리고 해당 객체가 자신이 예측한 객체가 맞을 확률

이 얼마나 되는지에 대한 확률에 대한 값을 나타낸다.

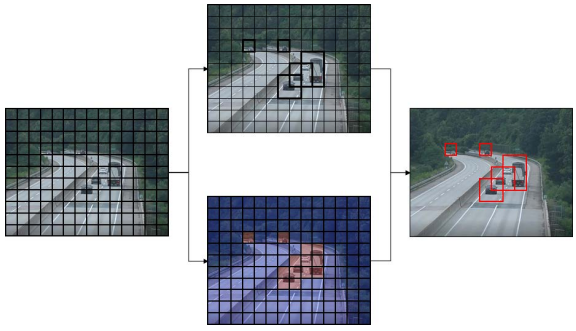


Fig. 3. YOLO detection system

따라서, 각 바운딩 박스는 클래스 확률을 갖게 되며 $x, y, w, h, score$ 로 구성되어 있다. (x, y)는 경계 상자의 중심값을 나타내는 상대적 위치 좌표를 나타내며, (w, h)는 너비와 높이의 상대 값을 나타낸다. 또한, 스코어는 Eq. (1)에 의해 해당 클래스의 확률 값이 계산되며, 이를 통해 7×7×30 텐서를 생성하여 학습한다(Redmon and Farhadi, 2017; Yu *et al*., 2019; Redmon *et al*., 2016).

$$Pr(Class_i|Object) \times Pr(Object) \times IoU_{pred}^{truth} = Pr(Class_i) \times IoU_{pred}^{truth} \quad (1)$$

where, Pr represents a probability, $Class$ represents a classification category and IoU represents a intersection over union.

이러한 YOLO V3는 IoU 0.5에서 mAP(mean Average Precision)가 57.9로 탐지 모델 중 하나인 RetinaNet과 동등하지만 속도가 약 4배 빠르므로 실시간 객체 탐지에 효과적이다(Redmon and Farhadi, 2017).

또한, 선행 연구에 따르면 Tung *et al*.(2019)의 다양한 조명 조건의 영상에서 YOLO를 이용하여 차량을 탐지하는 연구를 통해 TP (True Positive), TN (True Negative), FP (False Positive), FN (False Negative)을 각각 0.80, 0.99, 0.01, 0.20으로 높은 정확도를 도출하였다. Benjdira *et al*.(2019)는 무인 항공기를 이용한 차량 감지를 Faster R-CNN과 YOLO V3의 비교 연구를 수행하였다. Faster R-CNN의 경우 TP, FP, FN는 각각 578개, 2개, 150개, 재현율(recall)은 79.40%, F1 Score는 88.38%로 산정하였으며, YOLO V3의 경우 TP, FP, FN는 각각 751개, 2개, 7개이며, 재현율은 99.07%, F1 Score는 99.94%로 Faster R-CNN에 비해 높은 정확도를 도출하였다.

따라서, 본 연구에서는 선행 연구를 통해 차량 감지를 위한 정확도가 높은 YOLO V3 모델을 선정하였으며 COCO 데이터셋의 차량이 훈련된 YOLO V3 모델을 이용하여 실험을 수행하였다.

2.4 Affine transformation

부등각사상변환(affine transformation)은 등각사상변환의 축척인자에서 x, y 방향에 대해 각각의 축척인자가 다르다고 가정한 변환으로서 Eq. (2)와 같다(Park *et al*., 2009; Kim *et al*., 2008). 여기서, $a_0, a_1, a_2, b_0, b_1, b_2$ 는 부등각사상변환 매개변수, x', y'는 변환 전의 2차원 좌표를 나타내며, x, y는 변환 후의 좌표를 나타낸다.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

$$a_0 = x - axis\ translation$$

$$a_1 = Sx\ Cos(\alpha + \delta\alpha)$$

$$a_2 = -\ Sy\ Sin(\alpha) \quad (2)$$

$$b_0 = y - axis\ translation$$

$$b_1 = Sy\ Sin(\alpha + \delta\alpha)$$

$$b_2 = Sy\ Cos(\alpha)$$

where, S_x represents the scale factor along the $x-axis$, S_y represents the scale factor along the $y-axis$ and $\delta\alpha$ represents the non-orthogonality angle.

부등각사상변환의 매개변수를 결정하기 위한 관측방정식은 Eq. (3)과 같으며, 관측오차가 없다는 가정하에 수행하였다. 여기서, B는 매개변수 행렬, $\triangle$는 매개변수 $\triangle a_0, \triangle a_1, \triangle a_2, \triangle b_0, \triangle b_1, \triangle b_2$ 를 나타내며, f는 조건벡터이다.

$$B\triangle = f \quad (3)$$

따라서, 매개변수 $\triangle$는 Eq. (4)와 같이 계산할 수 있다.

$$\triangle = N^{-1}t \quad (4)$$

where, N represents B^TB and t represents B^Tf.

3. 실 험

3.1 실험 데이터 구축

본 연구의 대상지는 강원도 홍천군 동면 방량리에 위치한 중앙고속도로 일대이며 Fig. 4와 같이 스테레오 CCTV를 통해 영상을 취득하였다. 각 카메라별로 약 4분 30초 영상을 촬영하였으며, 차량을 연속적으로 탐지하기 위해 1초에 2개의 프레임으로 나누어 총 535장의 실험 데이터를 구축하였다.

Fig. 4. CCTV

또한, 다양한 환경에서 실험을 수행하기 위해 Fig. 5, Fig. 6과 같이 평상시와 안개가 발생할 경우의 영상을 취득하였으며 (a)는 좌측 카메라 영상, (b)는 우측 카메라 영상을 나타낸다. 두 영상의 촬영 시기 및 카메라 정보는 Table 2와 같다.

(a) Left camera

(b) Right camera

Fig. 5. Images in normal environments

(a) Left camera

(b) Right camera

Fig. 6. Images in foggy environments

3.2 Vehicle detection

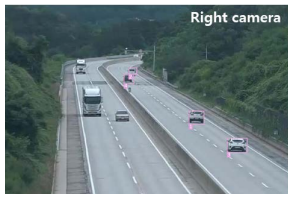
3.2.1 Vehicle detection in normal environments

Fig. 7은 평상시에 탐지한 차량의 일부이며 Fig. 7(a)는 좌측 카메라의 영상, Fig. 7(b)는 우측 카메라의 영상을 나타낸다. 좌측 카메라에서 탐지된 차량의 수는 6대이며, 우측 카메라에서 탐지된 차량의 수는 4대로 동일한 시기와 동일한 구간을 촬영하였음에도 불구하고 두 카메라에서 촬영된 영상의 특징이 서로 같지 않기 때문에 탐지된 차량과 차량 대수가 다른 것을 알 수 있었다. 또한, 본 연구의 데이터는 고속도로에 설치된 스테레오 CCTV로 도로 구간만 촬영을 수행하였기 때문에 차량에 대해서만 탐지가 되었으며 다른 객체로 오분류 되거나 다른 객체가 차량으로 탐지되는 경우는 발생하지 않았다.

Table 2. CCTV collected period and specification

Category		Normal environment	Foggy environment
CCTV collected period	Date	2018.07.05	2018.07.12
	Time(hour)	07	18
CCTV specification	Model	SNZ - 6320	
	Focal length	4.44mm ~ 142.6mm	
	Image size	640(width) × 480(height)	
	Baseline	45cm	

(a) Detected vehicles from the left camera



(b) Detected vehicles from the right camera

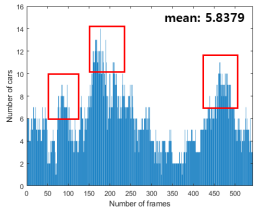
Fig. 7. Vehicle detection in normal environments

평상시에 촬영된 스테레오 CCTV 영상을 통해 탐지된 전체 차량의 수와 평균 탐지 정확도는 기구축된 COCO 데이터셋을 훈련한 모델로 실험을 통해 산출하였으나 이는 한국의 차량과 크기, 디자인 등에 차이가 있으므로 Table 3과 같은 결과를 도출하였다. 좌우 영상에서 탐지된 차량의 대수 및 정확도 차이는 각각 25대, 0.6%가 발생하는 것을 알 수 있었다.

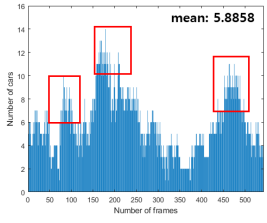
Table 3. Total number of detected vehicles in normal environments

Cameras	Number of vehicles	mean ACC(%)
Left camera	3,170	84.004
Right camera	3,195	83.432

평상시의 프레임별 추정한 차량 대수를 그래프로 나타나면 Fig. 8과 같으며 Fig. 8(a)는 좌측 영상에서 프레임별로 탐지한 차량 대수이며, Fig. 8(b)는 우측 영상에서 프레임별로 탐지한 차량 대수를 나타낸다. 좌측 영상을 이용한 평균 탐지 차량 대수는 5.8379대, 우측 영상을 이용한 평균 탐지 차량 대수는 5.8858대로 차이가 발생하였다. 이러한 차이를 보정하여 좌우 영상에서 동일한 차량을 탐지하여 교통량 산정의 정확도를 높이는 것이 필요하며 이를 위해 본 연구에서는 부등각사상변환을 이용하였다.



(a) Left camera



(b) Right camera

Fig. 8. Graph per frame in normal environments

3.2.2 Vehicle detection in foggy environments

Fig. 9는 안개가 발생한 경우 차량을 탐지한 결과 중 일부를 나타낸 것이다. 좌측 카메라에서 탐지된 차량의 수는 4대이며, 우측 카메라에서 탐지된 차량의 수는 5대를 탐지하였다. 또한, 안개가 발생한 경우는 차량의 형상이 안개에 가려져 평상시보다 비교적 짧은 구간의 차량만 탐지할 수 있었으며, 평상시와 동일하게 차량에 대해서만 탐지가 되었으며 다른 객체로 오분류 되거나 다른 객체가 차량으로 탐지되는 경우는 발생하지 않았다.

(a) Detected vehicles from the left camera

(b) Detected vehicles from the right camera

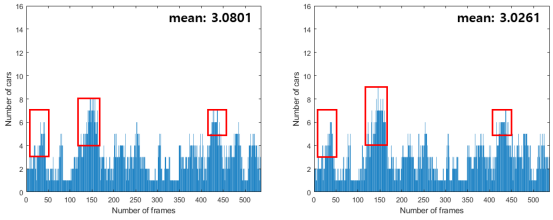
Fig. 9. Vehicle detection in foggy environments

안개가 발생한 경우 촬영된 좌우 영상을 통해 탐지된 전체 차량의 수와 평균 탐지 정확도는 Table 4와 같다. 안개가 발생한 경우 전체적으로 흰색 차량이 안개의 색상과 유사하여 평상시 보다 비교적 적게 탐 지되었으며, 좌우 영상에서 탐지된 차량의 대수 및 정확도 차이는 각각 31대, 0.8%가 발생하는 것을 알 수 있었다.

Table 4. Total number of detected vehicles in foggy environments

Cameras	Number of vehicles	mean ACC(%)
Left camera	1,669	84.142
Right camera	1,638	83.397

안개가 발생한 경우 영상에서 프레임별로 추정한 차량 대수를 그래프로 나타나면 Fig. 10과 같다. 좌측 영상에서 평균 탐지 차량 대수는 3.0801대, 우측 영상을 이용한 평균 탐지 차량 대수는 3.0261대로 나타났으며, 평상시보다 좌우 프레임별로 탐지한 차량의 수의 차이가 더 발생한 것을 확인할 수 있었다.



(a) Left camera (b) Right camera

Fig. 10. Graph per frame in foggy environments

3.3 Estimation of affine transformation parameters

단일 카메라 CCTV를 이용할 경우 영상의 모든 차량 객체를 추정하지 못할 뿐만 아니라 탐지되지 않는 차량이 상당수 존재하게 된다. 또한, 안개가 발생한 경우와 같이 도로상의 기상상태에 따라 차량 탐지 구간도 다르며 탐지된 차량과 탐지되지 않은 차량 수가 정확하지 않으므로 정확한 교통량 산정이 어렵게 된다. 따라서, 스테레오 CCTV 영상을 통해 좌우 영상 간의 부등각사상변환 매개변수를 추정하여 각 영상에서 탐지하지 못한 차량을 보완할 필요가 있다.

부등각사상변환 매개변수를 추정하기 위한 공액점은 Fig. 11, Table 5와 같으며, 평상시와 안개가 발생한 경우의 영상은 동일한 카메라로 촬영하였기 때문에 공액점은 같다.

(a) Left camera (b) Right camera

Fig. 11. Conjugate points

Table 5. Conjugate points

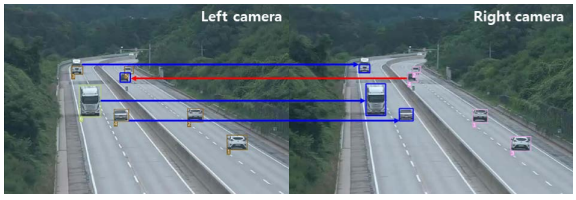
ID	Left Camera		Right Camera	
	X	Y	X	Y
1	800.525	795.528	800.426	799.445
2	744.475	382.459	768.492	382.551
3	139.497	805.425	136.541	795.415
4	337.556	481.423	353.167	474.731

공액점을 통해 추정한 좌측에서 우측, 우측에서 좌측 영상으로의 변환을 위한 부등각사상변환 매개변수와 사후 표준편차($\hat{\sigma_0}$) 값은 Table 6과 같다.

3.4 Applying an affine transformation

3.4.1 Applying an affine transformation to normal environments

평상시 기상상황의 좌우 영상에서 YOLO V3를 이용하여 탐지한 차량의 정확도는 평균 83.718%로 나타났다. 탐지 정확도에 근거해서 탐지된 객체는 높은 정확도로 차량 클래스로 분류되었기 때문에 참값이라고 가정하였다. Fig. 12(a)는 우측 영상에서 탐지된 차량을 부등각사상변환을 적용하여 좌측 영상에서 추가적으로 차량을 탐지한 결과를 나타낸다. 또한, Fig. 12(b)는 좌측 영상에서 탐지된 차량을 부등각사상변환을 적용하여 우측 영상에서 추가적으로 차량을 탐지한 결과를 나타내며, 부등각사상변환을 적용하여 하나의 차량에 여러 개의 바운딩 박스가 존재할 경우가 발생할 수 있으므로 바운딩 박스의 중심점을 계산하여 부등각사상변환의 사후 표준편차값의 2배 이내에 존재할 경우 중복된 바운딩 박스를 제외하였다. YOLO V3를 통해 탐지한 원본 영상은 Fig. 7에 나타낸 것과 같이 좌측 영상은 6대, 우측 영상은 4대를 각각 추정하였으나 좌우 영상 간의 변환관계를 고려한 경우 추정하지 못한 차량 객체도 추출함으로서 좌우 영상에서 동일한 7대를 탐지한 것을 확인할 수 있었다.



(a) Left camera (b) Right camera

Fig. 12. Applying an affine transformation to normal environments

평상시 기상환경에서 프레임별로 탐지한 차량 대수를 그래프로 나타나면 Fig. 13과 같으며, 영상 간의 변환관계를 고려하지 않았을 경우 Fig. 8에 나타낸 것과 같이 좌측은 평균

Table 6. Estimation of affine transformation parameters

Direction	X_T	Y_T	S_x	S_y	α	$\delta\alpha$	$\hat{\sigma_0}$
Left to Right	42.333	-16.458	1.005	1.007	0.057	-0.037	0.8066
Right to Left	-41.169	17.210	0.995	0.996	-0.057	0.037	0.7878

5.8379대, 우측은 평균 5.8858대의 차량을 탐지하였지만 좌우영상 간의 변환관계를 고려한 경우 평균 6.3421대의 차량을 탐지할 수 있었다.

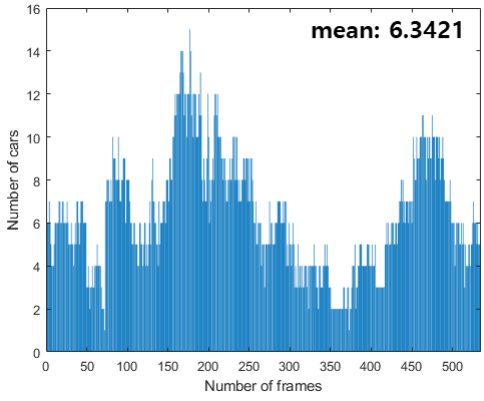


Fig. 13. Graph per frame with affine transformation parameters applied to normal environments

단일 카메라 CCTV일 경우와 좌우 영상 간의 변환관계를 고려한 스테레오 CCTV일 경우의 차량 탐지 수를 비교하면 Table 7과 같다. 좌우 영상 간의 변환관계를 고려한 경우 좌측 카메라의 경우 214대의 차량을 추가로 탐지할 수 있었으며 약 6.75%가 증가하였다. 또한, 우측 카메라일 경우 189대의 차량을 추가로 탐지할 수 있었으며 약 5.92%가 증가하는 것을 확인할 수 있었다.

Table 7. Comparison of detected vehicles in normal environments

Category	Left camera	Right camera
Number of vehicles	3,170	3,195
Applying affine transformation parameters	3,384	
Increased number of vehicles (proportion, %)	214(6.75%)	189(5.92%)

3.4.2 Applying an affine transformation to foggy environments

도로상에 안개가 발생한 기상상황의 좌우 영상에서 탐지한 차량의 정확도는 평균 83.770%이며 부등각사상변환 매개변수를 적용한 영상의 일부는 Fig. 14와 같다. Fig. 14(a)는 우측 영상에서 탐지된 차량을 부등각사상변환을 적용하여 좌측 영상에서 추가적으로 차량을 탐지한 결과를 나타내고 Fig. 14(b)는 좌측 영상에서 탐지된 차량을 부등각사상변환을 적용하여 우측 영상에서 추가적으로 차량을 탐지한 결과를 나타내며, 평상시와 동일하게 바운딩 박스의 중심점을 계산하여 부등각사상변환의 사후 표준편차 값의 2배 이내에 존재할 경우 중복된 바운딩 박스를 제외하였다. YOLO V3를 통해 차량을 탐지한 원본 영상은 Fig. 9와 같이 좌측은 4대, 우측은 5대를 추정하였으나 좌우 영상 간의 변환관계를 고려한 경우 좌우 영상에서 동일한 6대의 차량을 탐지한 것을 확인할 수 있었다.

(a) Left camera **(b) Right camera**

Fig. 14. Applying an affine transformation to foggy environments

안개가 있는 영상에서 프레임별로 탐지한 차량 대수를 그래프로 나타내면 Fig. 15와 같다. 좌우 영상 간의 변환관계를 고려하지 않은 경우 좌측 영상의 평균 차량 탐지대수는 Fig. 10(a)와 같이 3.0821대, 우측 영상의 평균 차량 탐지대수는 Fig. 10(b)와 같이 3.0261대이다. 이에 반해, 좌우 영상 간의 변환관계를 고려한 경우 평균 차량 탐지대수는 3.4467대로 증가한 것을 확인할 수 있었다.

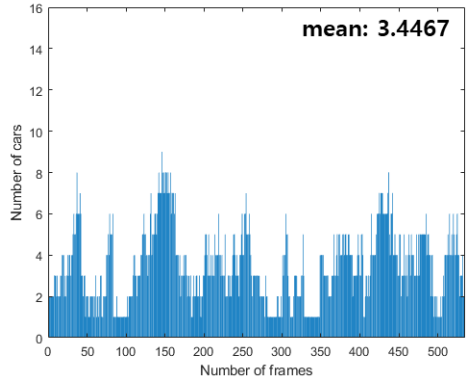


Fig. 15. Graph per frame with affine transformation parameters applied to foggy environments

단일 카메라 CCTV일 경우와 좌우 영상 간의 변환 관계를 고려한 스테레오 CCTV일 경우에 탐지된 차량 수를 비교하면 Table 8과 같다. 좌우 영상 간 변환관계를 고려한 경우 좌측 카메라에서 180대의 차량을 추가로 탐지할 수 있었으며 약 10.785%가 증가하였다. 또한, 우측 카메라에서 211대의 차량

을 추가로 탐지할 수 있었으며 약 12.882%가 증가하는 것을 확인할 수 있었다.

Table 8. Comparison of detected vehicles in foggy environments

Category	Left camera	Right camera
Number of vehicles	1,669	1,638
Applying affine transformation parameters	1,849	
Increased number of vehicles (proportion, %)	180 (10.785%)	211 (12.882%)

3.5 Estimation of traffic volume

도로 교통량이란 단위 시간(일, 시간, 15분, 5분 등)에 어떤 지점을 통과하는 차량의 대수를 말한다(Huh *et al*., 2013).

평상시 스테레오 CCTV 영상에서 부등각사상변환 매개변수를 고려하여 교통량을 계산하면 초당 약 6.3421대가 통과한다. 이를 단위 시간을 고려하여 환산하여 5분, 15분, 1시간, 1일 교통량으로 변환하면 Table 9와 같다.

Table 9에 나타낸 것과 같이 평상시 도로 환경에서 스테레오 CCTV 영상을 통해서 분석한 1일 교통량은 약 12,454대인데, 동일한 촬영날짜의 한국도로공사 고속도로 공공데이터 포털에서 제공하는 일별 합계 교통량은 13,380대로 나타나 약 926대(약 7%)의 오차가 발생하였다.

또한, 안개가 발생한 기상상황을 고려한 경우의 스테레오 CCTV 영상을 통해 분석한 교통량을 계산하면 초당 약 3.4467대가 통과하며, 1일 교통량은 약 6,768대로 계산되었다. 평상시와 유사하게 초당 교통량을 단위 시간을 고려하여 환산하여 5분, 15분, 1시간, 1일 교통량으로 변환하면 Table 10과 같다. Table 10에서 알 수 있듯이 안개가 발생한 경우의 교통량이 평상시보다 약 46% 줄어든 6,768대인 것을 알 수 있었다.

Table 10에 나타낸 것과 같이 안개가 발생한 기상상황일 경우 스테레오 CCTV 영상을 통해서 분석한 1일 교통량은 약 6,768대인데, 동일한 촬영날짜의 한국도로공사 고속도로 공공데이터 포털에서 제공하는 일별 합계 교통량은 9,051대로 나타나 약 2,283대(약 25%)의 오차가 발생하였다. 이는 안개가 발생한 약 4분 30초 스테레오 CCTV 영상을 통해 1일 교통량을 산정하였기 때문으로 1일 동안 촬영된 영상을 이용하여 교통량을 산정할 필요가 있는 것으로 나타났다.

4. 결 론

본 연구는 단일 CCTV 영상에서 딥러닝을 이용하여 모든 객체를 탐지하기 어렵기 때문에 스테레오 CCTV 영상에서 딥러닝을 이용하여 객체를 탐지한 후, 각 영상에서 탐지하지 못한 차량에 대해 부등각사상변환을 적용하여 영상 내 모든 차량을 탐지하고 교통량을 산정하는 것으로 다음과 같은 결론을 도출하였다.

첫째, 현재 교통량 산정은 AVC, VDS, TCS 등의 추가적인 장비 설치가 필요하며 CCTV를 통한 인력식 조사를 병행하여 교통량을 산정하고 있기 때문에, 하드웨어 설치비용과 인력 운용에 많은 비용이 소요되고 있다. 이에 본 연구에서는 스테레오 CCTV 영상과 딥러닝 기술을 접목하여 도로상의 교통량 산정을 자동화 하기 위한 방안을 제시하였다.

둘째, 평상시와 안개가 발생한 기상상황의 경우에 촬영된 스테레오 CCTV 영상에 대해 차량을 학습한 데이터를 통해 사람의 수작업이 아닌 자동으로 차량을 탐지할 수 있었다.

셋째, 단일 카메라 CCTV 영상의 경우 영상에서 모든 차량 객체를 탐지하는데 어려움이 발생하였으나, 스테레오 CCTV 영상의 부등각사상변환 매개변수를 추정하여 적용하였을 경우 단일 카메라 CCTV 영상에서 탐지하지 못한 차량도 추가적으로 탐지할 수 있었으며 이를 통해 교통량 산정의 정확도를 높일 수 있었다.

본 연구에서는 스테레오 CCTV의 일부 영상만 활용하였으나 촬영된 전체 영상에 대해 딥러닝을 적용하여 교통량을 산정할 경우 도로 구간의 시간대 별 교통량을 추정할 수 있

Table 9. Estimation of traffic volume in normal environments

	5 minute	15 minute	1 hour	1 day
Number of vehicles	43.242	129.725	518.899	12,453.578

Table 10. Estimation of traffic volume in foggy environments

	5 minute	15 minute	1 hour	1 day
Number of vehicles	23.502	70.501	282.003	6,768.065

을 것으로 사료된다. 또한, 다양한 촬영 시기와 도로상의 기상환경이 다른 경우의 차량 데이터를 학습하여 CCTV 영상에 촬영된 차량의 종류를 세분화하여 탐지를 수행할 경우 차량의 종류를 구분하고 차량별 교통량을 산정할 수 있을 것으로 사료된다.

References

Ammour, N., Alhichri, H., Bazi, Y., Benjdira, B., Alajlan, N., and Zuair, M. (2017), Deep learning approach for car detection in UAV imagery. *Remote Sensing*, Vol. 9, No. 4, pp. 312-326.

Anindra, F., Soeparno, H., and Napitupulu, T. A. (2018), CCTV traffic congestion analysis at pejompongan using case based reasoning. In *2018 International Conference on Information and Communications Technology (ICOIACT)*, 6-7 March, Yogyakarta, Indonesia, pp. 861-865.

Barthélemy, J., Verstaevel, N., Forehead, H., and Perez, P. (2019), Edge-computing video analytics for real-time traffic monitoring in a smart city, *Sensors*, Vol. 19, No. 9, pp. 2048-2078.

Benjdira, B., Khursheed, T., Koubaa, A., Ammar, A., and Ouni, K. (2019), Car detection using unmanned aerial vehicles: Comparison between faster r-cnn and yolov3. In *2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)*, IEEE, 5-7 February, Muscat, Oman, pp. 1-6.

Choi, I.K. and Yoo, J.S. (2017), Object detection in road environment CCTV images using deep learning. *The Institute of Electronics and Information Engineers*, 24-25 November, Incheon, Korea, pp. 627-629.

Du, X., Ang, M.H., and Rus, D. (2017), Car detection for autonomous vehicle: LIDAR and vision fusion approach through deep learning framework. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 24-28 September, Vancouver, Canada, pp. 749-754.

Han, S.H., Shin, Y.S., and Lee, J.Y. (2019), A study on the evaluation technique of intelligent security technology based on spatial information : multi-CCTV collaboration technology. *The Journal of the Korea Academia-Industrial cooperation Society*, Vol. 20, No. 7, pp. 111–118. (in Korean with English abstract)

Hong, G.S., Eom, T.J., and Kim, B.G. (2011), Development of vision-based monitering system technology for traffic analysis and surveillance. *Jouranl of Information and Security*, Vol. 11, No. 4, pp. 59-66.

Huh, M.H., Shin, S.Y., and Lee, Y.W. (2013), Traffic measurement : moving vehicle method using CCTV. *Journal of the Korea Institute of Information and Communication Engineering*, Vol. 17, No. 11, pp. 2575-2580. (in Korean with English abstract)

Jeong, D.H. and Jeong, W.T. (2019), Prediction of rolling noise based on machine learning technique using rail surface roughness data, *Journal of the Korean Society for Railway*, Vol. 22. No. 3, pp. 209-217. (in Korean with English abstract)

Jo, S.H., Kim, C.G., Lim, H.Y., and Shin, Y.T. (2018), A study on the traffic flow analysis method based on change detection for traffic video data. *Journal of Information Technology and Architecture*, Vol. 15, No. 3, pp. 373–382. (in Korean with English abstract)

Kim, J.H. and Choi, D.H. (2019), Implementation of a vehicle traffic and speed estimation system using faster R-CNN. *The Journal of Korean Institute of Communications and Information Sciences*, Vol. 44, No. 9, pp. 1754-1758. (in Korean with English abstract)

Kim, S.S., Jung, J.H., Kim, E.M., Yoo, H.H., and Sohn, H.G. (2008), Geocoding of low altitude UAV imagery using affine transformation model. *Journal of Korean Society for Geospatial Information Science*, Vol. 16, No. 4, pp. 79–87. (in Korean with English abstract)

Kim, Y.M., Lee, J.Y., Yoon, I.L., Han. T.J., and Kim, C.Y. (2018), CCTV object detection with background subtraction and convolutional neural network. *The Korean Institute of Information Scientists and Engineers*, Vol. 24, No. 3, pp. 151–156. (in Korean with English abstract)

Lee, G.W. and Yom, J.H. (2018), Design and implementation of web-based automatic preprocessing system of remote sensing imagery for machine learning modeling. *The Journal of Korean Society for Geospatial Information Science*, Vol. 26, No. 1, pp. 61–67. (in Korean with English abstract)

Lee, T.H., Kim, K.J., Yun, K.S., Kim, K.J., and Choi, D.H.

(2018), A method of counting vehicle and pedestrian using deep learning based on CCTV. *The Korean Institute of Intelligent Systems*, Vol. 28, No. 3, pp. 219–224. (in Korean with English abstract)

Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollar, P., and Zitnick, C.L. (2014), Microsoft coco: Common objects in context. In *European conference on computer vision*, Springer, Cham, 6-12 September, Zurich, Switzerland, pp. 740-755.

Mundhenk, T. N., Konjevod, G., Sakla, W. A., and Boakye, K. (2016), A large contextual dataset for classification, detection and counting of cars with deep learning. *In European conference on computer vision*, Springer, Cham, 8-16 October, Amsterdam, Netherlands, pp. 785-800.

Park, G.M. and Bae, Y.C. (2019), Performance comparison of machine learning in the various kind of prediction. *The Journal of the Korea Institute of Electronic Communication Sciences*, Vol. 14, No. 1, pp. 169–178. (in Korean with English abstract)

Park, S.Y., Lee, J.B., Park, Y.J., and Yu, K.Y. (2009), The study on coordinate transformation for updating of digital map from construction drawing data. *Journal of Korean Society of Surveying, Geodesy*, *Photogrammetry and Cartography*, Vol. 27, No. 2, pp. 281–288. (in Korean with English abstract)

Peppa, M.V., Bell, D., Komar, T., and Xiao, W. (2018), Urban traffic flow analysis based on deep learning car detection from CCTV image series. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, Vol. 42, No. 4, pp. 499-506.

Redmon, J. and Farhadi, A. (2017), YOLO9000 : Better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 21-26 July, Honolulu, USA, pp. 7263–7271.

Redmon, J. and Farhadi, A. (2018), YOLO V3: An incremental improvement, *arXiv preprint arXiv:1804.02767*, pp. 1-6.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016), You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 27-30 June, Las Vegas, USA, pp. 779–788.

Sirirattanapol, C., Nagai, M., Witayangkurn, A., Pravinvongvuth, S., and Ekpanyapong, M. (2019), Bangkok CCTV image through a road environment extraction system using multi-label convolutional neural network classification. *ISPRS International Journal of Geo-Information*, Vol. 8, No. 3, pp. 128-143.

Traffic Monitoring System. (2018), Road traffic investigation, *Ministry of Land, Infrastructure and Transport*, URL: http://www.road.re.kr/(last date accessed: 22 December 2019).

Tung, C., Kelleher, M.R., Schlueter, R.J., Xu, B., Lu, Y.H., Thiruvathukal, G.K., Chen, Y.K., and Lu, Y. (2019), Large-scale object detection of images from network cameras in variable ambient lighting conditions. In *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, IEEE, 28-30 March, California, USA, pp. 393-398.

Xu, Y., Wu, L., Xie, Z., and Chen, Z. (2018), Building extraction in very high resolution remote sensing imagery using deep learning and guided filters. *Remote Sensing*, Vol. 10, No. 1, pp. 144-161.

Young, T., Hazarika, D., Poria, S., and Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational IntelligenCe Magazine*, Vol. 13, No. 3, pp. 55-75.

Yu, J.H., Han, Y.J., and Hahn, H.S. (2019), Improving performance of YOLO network using multi-layer overlapped windows for detecting correct position of small dense objects. *Journal of The Korea Society of Computer and Information*, Vol. 24, No. 3, pp. 19–27.