

차선 인식에서의 효과적인 자기 지도 학습 전략

이현종, 강윤석, 조우성, 윤의현, 이재구*

*국민대학교

*jaekoo@kookmin.ac.kr

Self-Supervised Learning Strategies for Lane Detection

Hyunjong Lee, Yunseok Kang, Wooseong Cho, Euihyun Yoon, Jaekoo Lee*
College of Computer Science, *Kookmin University.

요약

일반적으로 자기 지도 학습(Self-Supervised Learning) 후 전이 학습(Transfer Learning)을 하는 방식은 컴퓨터 비전 과업에서 성능 향상에 효과적인 전략으로 알려져 있다. 주로 일반적인 다양한 객체를 가지고 있는 ImageNet 데이터 집합이나 전이할 과업과 같은 영역(Domain)의 데이터 집합을 자기 지도 학습에 사용한다. 하지만 차선 인식 분야에서는 이에 대한 연구가 다른 분야에 비해 미비하다. 따라서 본 논문에서는 차선 인식에서 효과적인 자기 지도 학습 전략에 대해 연구하였다. 다른 분야와 유사하게 ImageNet 과 차선 인식 데이터 집합들을 활용하여 자기 지도 학습을 수행하였다. 이후 차선 인식 데이터 집합인 CULane 데이터 집합에서 전이 학습을 하여 결과를 비교 분석했다. 결과적으로 차선 인식에서 ImageNet 이 차선 영역의 데이터 집합보다 자기 지도 학습에서 효과적임을 실험적으로 확인했다.

I. 서론

컴퓨터 비전 분야에서의 자기 지도 학습은 객체 탐지 분야, 의료 분야, 위성 분야 등 다양한 응용 분야에서 활발히 연구가 이루어지고 있다[1]. 주로 자연 이미지를 지니고 있는 ImageNet[2]이나 전이할 과업과 유사한 영역의 데이터 집합을 사용하여 자기 지도 학습을 수행하는 방법이 있다. 하지만 차선 인식 분야에서는 자기 지도 학습에 대한 연구가 상대적으로 부족한 상태이다. 특히 차선의 영역에서 자기 지도 학습을 진행하는 연구는 거의 이루어지지 않았다.

그림 1(a, b)을 보면 ImageNet 은 다양한 객체를 포함하고 각 이미지가 서로 다른 특징을 지니지만, 차선 영역 데이터는 하늘과 도로의 형태가 유사하며, 도로 위에서 데이터가 수집되기 때문에 등장하는 객체도 제한이 있다. 이러한 차이점이 존재하기 때문에 차선 인식 과업에서 효과적인 자기 지도 학습 전략에 대한 연구가 필요하다. 따라서 본 논문에서는 ImageNet 과 차선 영역 데이터 집합을 비교 분석하여 차선 인식에서의 자기 지도 학습 전략에 대한 연구를 수행하였다.

II. 본론

그림 1 은 실험 개요이다. 일반적인 전이 학습은 그림 1(a, f)에서 보이듯 ImageNet 을 사용하여 지도 학습을 수행한 이후 차선 데이터 집합을 사용하여 전이 학습을 수행한다. 하지만 최근 그림 1(b, f)에서 보이듯 지도 학습이 아닌 자기 지도 학습을 사용하여 전이 학습을 수행하는 연구가 성능 측면에서 우세를 보이고 있다. 따라서 본 논문에서는 차선 인식에서 효과적인 자기 지도 학습 전략을 세우기 위해 그림 1(b, c)에서 보이듯 ImageNet 데이터 집합과 차선 데이터 집합을 각각 사용하여 자기 지도 학습을 한 이후 차선 데이터 집합에 전이 학습을 수행하였다.

이때 자기 지도 학습은 정답값 없이 이미지의 일반적인 특징을 학습하기 때문에 많은 데이터를 요구한다. 하지만 ImageNet 과 달리 CULane 은 데이터

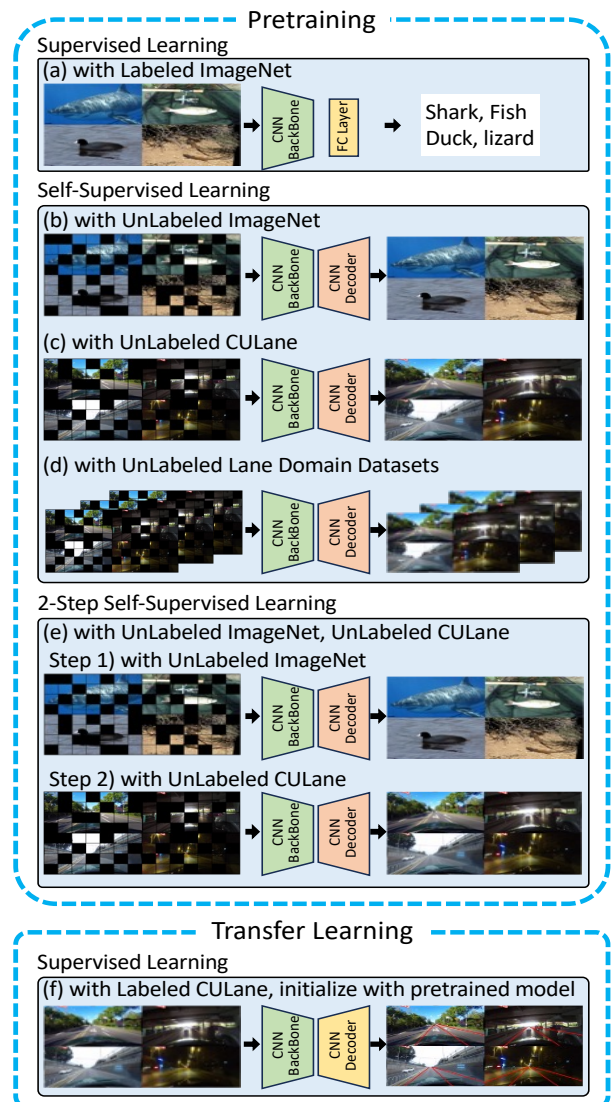


그림 1. 실험 개요

표 1. CULane 에서의 RESA 실험 결과

Strategies	F1 Score	Total	Normal	Crowd	Night	No line	Shadow	Arrow	Dazzle	Curve
Supervised with ImageNet		75.275	92.167	73.195	70.149	48.083	73.033	88.196	68.847	71.195
Self-Supervised with ImageNet		75.509	92.262	73.818	70.313	47.432	72.184	88.301	67.410	70.089
Self-Supervised with CULane		73.967	91.837	72.749	67.369	46.017	74.035	87.296	63.955	66.892
Self-Supervised with Lane domain datasets		73.628	91.729	71.497	67.318	46.637	72.256	87.781	62.995	66.934
2-Step Self-Supervised with ImageNet, CULane		74.641	92.240	72.829	68.962	46.423	72.579	87.883	62.618	69.002

양이 현저히 적다. 따라서 본 논문에서는 이러한 문제를 해소하기 위해 그림 1(d, e)와 같은 실험을 추가로 진행하였다. 먼저 그림 1(d)는 CULane 의 부족한 데이터를 보충하기 위해 다른 차선 영역의 데이터를 추가하여 학습을 진행한 것이다. 하지만 이 경우 여러 도메인이 섞여 학습이 불안정하고, 성능이 오히려 감소하는 문제가 있었다. 따라서 이러한 문제점을 해소하기 위해 ImageNet 에서 자기 지도 학습을 먼저 수행한 이후 목표 영역에서 자기 지도 학습을 추가로 진행한 선행 연구[4]처럼, CULane 에서 추가로 학습을 수행하였다.

III. 실험 및 분석

본 논문에서는 ImageNet, CULane 데이터 집합, 그리고 차선 영역의 여러 데이터 집합을 혼합한 차선 영역 데이터 집합을 사용하여 자기 지도 학습을 수행했다. 이때 차선 영역 데이터 집합은 CULane 88,880 장, Tusimple 3,626 장, LLMAS 100,000 장, BDD100k 100,000 장을 합쳐 총 292,506 장의 데이터를 구성했다. 이는 기존 CULane 데이터 개수보다 약 3.3 배 많은 수치이다. 그리고 자기 지도 학습으로 현재 SOTA(State-Of-The-Art)를 달성한 MIM(Masked Image Modeling) 기반 모델인 Spark[5]를 사용하였다. 그리고 이미지 분할(Image Segmentation) 기반 모델인 RESA[6] 모델을 기준 삼아 F1 측정값을 수행하였다.

표 1 의 실험 결과를 보면 기존 ImageNet 으로 지도 학습을 수행한 방법에 비해 ImageNet 으로 자기 지도 학습을 한 경우 F1 측정값이 0.234 높게 평가되었다. 하지만 차선 영역에서 자기 지도 학습을 실험한 경우 지도 학습을 수행한 방법에 비해 F1 측정값이 모두 감소하였다. CULane 데이터 집합에서는 자기 지도 학습을 한 경우 지도 학습에서 수행한 방법에 비해 F1 측정값이 1.308 감소하였고 차선 영역 데이터 집합에서 자기 지도 학습을 한 경우 지도 학습에서 수행한 방법에 비해 F1 측정값이 1.647 감소하였다.

이러한 결과는 이미지 분할 작업에서는 다양한 특징이 학습되어야 한다는 기존 연구[7]와 경향성이 유사하다. 즉 차선 영역 데이터 집합은 도로 위에서 수집되기 때문에 유사한 형태와, 적은 객체를 지니게 되어 학습에 필요한 특징이 부족하다고 판단할 수 있다. 특히 이러한 점은 ImageNet 으로 자기 지도 학습을 한 것에 비해 2-Step 으로 자기 지도 학습을 한 경우 성능이 0.634 낮아지지만, CULane 데이터 집합으로 자기 지도 학습을 한 경우 2-Step 으로 자기 지도 학습을 한 경우보다 성능이 0.674 높아지는 점에서 확인할 수 있다.

그림 2 는 각각 ImageNet 에서 자기 지도 학습과 지도 학습을 한 이후 전이 학습을 수행한 결과이다. 그림에서 보듯이 입력 이미지에 차선이 잘 드러나지 않는 경우 자기 지도 학습이 더 좋은 결과를 보였다.

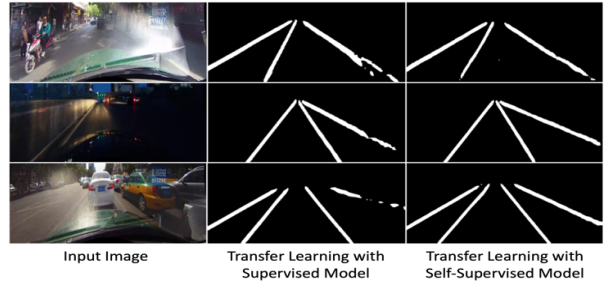


그림 2. 차선 인식 결과 시각화

IV. 결론

본 논문에서는 차선 인식에 효과적인 자기 지도 학습 전략에 대해 조사했다. 차선 인식의 경우 다른 컴퓨터 비전 분야와 달리 같은 영역의 데이터로 자기 지도 학습을 할 경우 성능이 하락했다. 또한 ImageNet 데이터 집합을 사용하여 자기 지도 학습을 하는 경우 효과적으로 다른 방식에 비해 성능을 높일 수 있는 전략임을 실험적으로 입증했다. 추후 ImageNet 데이터 집합에서 자기 지도 학습을 활용하여 차선 인식의 성능을 높이는 연구를 기대하는 바이다.

ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.RS-2023-00212484,복잡한 실제 주행환경에서 설명 가능한 움직임 예측).

본 연구는 2022년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음"(2022-0-00964)

참고 문헌

- [1] Shurrab, Saeed, and Rehab Duwairi. "Self-supervised learning methods and applications in medical imaging analysis: A survey." PeerJ Computer Science 8 (2022): e1045.
- [2] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. IJCV, 2015.
- [3] Pan, Xingang, et al. "Spatial as deep: Spatial cnn for traffic scene understanding." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 32. No. 1. 2018.
- [4] Shekoofeh, Azizi, et al. "Big self-supervised models advance medical image classification." arXiv preprint arXiv:2101.05224 (2021).
- [5] Tian, Keyu, et al. "Designing bert for convolutional networks: Sparse and hierarchical masked modeling." arXiv preprint arXiv:2301.03580 (2023).
- [6] Zheng, Tu, et al. "Resa: Recurrent feature-shift aggregator for lane detection." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 35. No. 4. 2021.
- [7] He, Kaiming, Ross Girshick, and Piotr Dollár. "Rethinking imagenet pre-training." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.