

위성 영상 데이터 집합 사이의 도메인 간극 일반화

하정민, 이재구*
국민대학교

*jaekoo@kookmin.ac.kr

Bridging the Domain Gap in Remote Sensing Image Datasets for Domain Generalization

Jungmin Ha, Jaekoo Lee*
*Kookmin Univ.

요약

본 논문은 위성 영상 인식 과업에서 자기 지도 학습을 위해 데이터 집합 간의 도메인 간극을 확인하고, 도메인 일반화 기법을 사용하여 도메인 간극 극복 가능성을 확인했다. 도메인 간극을 확인하기 위해 위성 영상 벤치마크 데이터 집합인 UCMerced Land Use, AID, Optimal 31, PatternNet, WHU RS-19를 사용했으며, 하나의 데이터 집합에 학습한 후 나머지 데이터 집합에 검증을 진행했다. 그 결과 데이터 집합 간의 도메인 간극을 확인했다. 또한, 도메인 일반화 기법인 RSC를 적용했을 때 F1 score는 최대 0.21, Accuracy는 최대 31.9 % 성능 향상을 확인하여 도메인 간극을 극복할 가능성을 확인했다.

I. 서론

위성 영상 데이터는 국토 관리, 재난 예방, 환경 오염 등을 분석하는 것에 사용된다. 또한, 최근에는 인공지능과 결합하여 위성 영상을 인식하는 연구가 진행되고 있으며, 많은 양의 위성 영상 벤치마크 데이터 집합들(Benchmark Datasets)이 구축되었다. 그러나 위성 영상 데이터는 수집 비용이 많이 들기 때문에 이미지 분류 과업뿐 아니라 객체 탐지, 객체 분할 과업에 사용할 수 있는 데이터의 수가 적다. 이런 점을 극복하기 위해 비지도 학습에서 사용하는 전이 학습을 활용할 수 있다. 그러나 비지도 학습을 수행하기 위해서는 먼저 정답이 없는 많은 양의 이미지가 필요하고 이미지 간의 도메인이 같아야 하는 특징이 있다.

따라서 본 논문에서는 먼저 위성 영상 데이터 사이의 도메인 간극(Domain Gap)이 존재하는지를 알아보았다. 실험에는 ResNet50[1]이 사용되었으며, 모든 데이터 집합에서 같은 구조의 네트워크로 학습하기 위해 모든 데이터 집합의 클래스(Class)들을 수로 네트워크의 출력 크기로 설정했다. 이후 한 데이터 집합에서 학습을 진행하고 다른 모든 데이터 집합에 검증을 진행했다. 검증 지표는 F1 score와 Accuracy를 사용했으며, 대다수 데이터 집합에서 중복으로 출현하는 클래스에 대한 검증을 진행했다. 그 결과 위성 영상 데이터 집합 사이의 도메인 간극을 확인하고, 나아가 도메인 일반화(Domain Generalization)방법인 Representation Self-Challenging(RSC)[2]을 사용하여 데이터 집합 간의 도메인 간극을 줄일 수 있음을 확인했다.

II. 본론

딥러닝의 특징은 인공 신경망을 사용하여 많은 데이터를 통해 각 클래스의 일반화된 복잡한 특징을 학습하는 것이다. 그러나 위성 영상 데이터는 사용 가능한 데이터의 수가 부족하므로 딥러닝에 적용하기 어려운 단점이 있다. 따라서 정답 값없이 일반적인 특징을 학습하고 적은 데이터에 전이하는 비지도 학습 방법이 적합하다. 비지도 학습을 활용한다면 여러 데이터 집합을 합쳐 하나의 거대한 데이터 집합을 학습에 활용할 수 있다.

그러나 데이터 집합 사이에 도메인 간극이 존재한다면 그림 1.과 같이 학습이 진행되는 동안 도메인 간극이 더 벌어지는 문제가 발생한다[3]. 따라서 도메인 간극을 줄이는 도메인 일반화 기법을 적용한다면 도메인이 다르더라도 데이터들을 병합하여 비지도 학습에 많은 데이터를 활용할 수 있을 것으로 기대할 수 있다. 본 논문에서 도메인 간극이 존재하는 데이터 집합 사이의 도메인 간극을 줄이기 위해 네트워크에 규제를 주어 일반적인 특징을 학

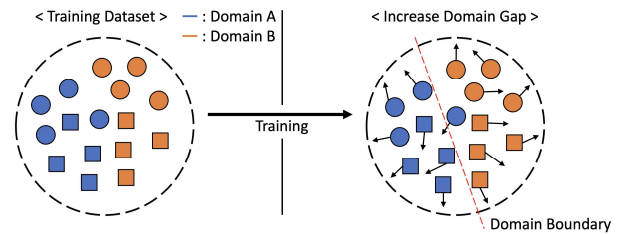


그림 1. 자기지도 학습에서 도메인 간극이 존재할 때 데이터 분포 학습하는 RSC를 사용했다. RSC는 그림 2.와 같이 특정 클래스 혹은 도메인에 편향되어 학습되는 것을 방지하기 위해 정답 값으로부터 기울기를 먼저 계산하여 분류기를 통과하기 전 특징맵에 전달되는 기울기를 구한다. 이후 값이 크게 나온 특징맵의 일부분을 마스킹을 적용한 후 분류기에 통과한다. 이런 과정으로 분류를 학습하면 규제를 통해 특정 도메인에 편향되는 것을 방지할 수 있다.

III. 실험

본 논문에서는 위성 데이터 집합 사이의 도메인 간극을 알아보고 도메인 간극이 도메인 일반화 기법을 통해 완화될 수 있는지 알아보기 위해 실험을 두 가지 진행했다. 먼저 도메인 간극을 알아보기 위해서 ResNet50을 사용하여 원천 도메인(Source Domain)에 학습을 진행한 후 목표 도메인(Target Domain)의 중복 클래스에 대한 성능을 측정했다. 그 후 RSC기법을 적용하여 네트워크를 원천 도메인에 학습한 후 목표 도메인에 검증하는 것으로 도메인 간극이 완화되는 것을 확인했다. 측정 지표는 F1 score와 Accuracy를 사용했으며, 모든 데이터 집합에 공통으로 포함된 클래스에 대해 측정했다. 공통된 클래스는 총 3가지로 Airport, Beach, Baseball Filed가 있다.

표 1.에서 각 데이터 집합에서 학습한 후 모든 데이터 집합에 대해 검증을 진행한 결과를 확인할 수 있다. 자신의 데이터에 대해 검증을 진행할

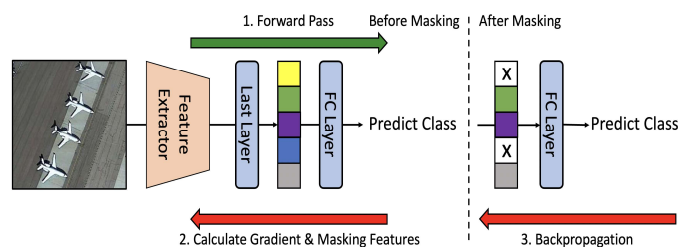


그림 2. RSC의 학습 과정, X는 마스킹된 특징맵

표 1. 데이터 집합 별 F1 score 성능 측정 및 성능 등락

Train data \ Test data	UCMerced Land Use		AID		Optimal 31		PatternNet		WHU RS-19	
	F1 score	Diff.	F1 score	Diff.	F1 score	Diff.	F1 score	Diff.	F1 score	Diff.
UC Merced Land Use	0.92 (0.95)	- 0.03	0.50 (0.49)	+ 0.01	0.57 (0.42)	+ 0.15	0.65 (0.61)	+ 0.04	0.30 (0.27)	+ 0.03
AID	0.35 (0.34)	+ 0.01	0.88 (0.91)	- 0.03	0.53 (0.48)	+ 0.05	0.22 (0.26)	+ 0.04	0.41 (0.37)	+ 0.04
Optimal 31	0.62 (0.44)	+ 0.18	0.71 (0.68)	+ 0.03	0.94 (0.86)	+ 0.08	0.66 (0.74)	- 0.08	0.31 (0.10)	+ 0.21
PatternNet	0.62 (0.43)	+ 0.19	0.52 (0.51)	+ 0.01	0.71 (0.70)	+ 0.01	1.00 (0.99)	+ 0.01	0.24 (0.04)	+ 0.20
WHU RS-19	0.42 (0.32)	+ 0.10	0.95 (0.95)	0.00	0.53 (0.28)	+ 0.25	0.07 (0.02)	+ 0.05	0.93 (0.75)	+ 0.18

표 2. 데이터 집합 별 Accuracy 성능 측정 및 성능 등락

Train data \ Test data	UCMerced Land Use		AID		Optimal 31		PatternNet		WHU RS-19	
	Accuracy	Diff.	Accuracy	Diff.	Accuracy	Diff.	Accuracy	Diff.	Accuracy	Diff.
UC Merced Land Use	89.0 % (90.9 %)	- 1.9 %	47.3 % (40.3 %)	+ 7.0 %	51.3 % (37.7 %)	+ 13.6 %	50.3 % (48.3 %)	+ 2.0 %	41.0 % (28.5 %)	+ 12.5 %
AID	29.4 % (26.2 %)	+ 3.2 %	80.8 % (85.9 %)	- 5.1 %	43.9 % (36.3 %)	+ 7.6 %	13.9 % (16.9 %)	- 3.0 %	41.3 % (37.1 %)	+ 4.2 %
Optimal 31	51.1 % (33.4 %)	+ 17.7 %	58.3 % (55.5 %)	+ 2.8 %	87.7 % (75.3 %)	+ 12.4 %	57.2 % (61.7 %)	- 4.5 %	35.9 % (9.2 %)	+ 26.7 %
PatternNet	53.9 % (42.1 %)	+ 11.8 %	36.9 % (38.9 %)	- 2.0 %	59.8 % (60.3 %)	- 0.5 %	99.6 % (99.0 %)	+ 0.6 %	34.9 % (3.0 %)	+ 31.9 %
WHU RS-19	37.1 % (23.8 %)	+ 13.3 %	90.5 % (91.5 %)	- 1.0 %p	49.5 % (21.0 %)	+ 28.5 %	3.8 % (1.0 %)	+ 2.8 %	84.3 % (63.3 %)	+ 21.0 %

때는 학습 데이터와 검증 데이터를 8:2의 비율로 나눠 사용했으며, 표 1.에서 대각선 방향에 해당하는 지표를 각 데이터 집합의 기준성능으로 설정했다. 해당 지표를 기준으로 다른 데이터 집합에 대한 성능을 비교했을 때, AID에 학습한 후 WHU RS-19에 검증을 진행한 것 외에는 모두 성능이 하락한 것을 확인할 수 있다. 이는 데이터 집합 사이에 도메인 간극이 존재함을 알 수 있는 결과이다. 또한, AID에 학습한 후 WHU RS-19에 검증한 결과가 높게 나온 것은 데이터의 도메인이 유사하고 WHU RS-19가 AID에 비해 단순한 데이터 분포를 하고 있기 때문으로 해석할 수 있다.

일반적으로 도메인 간극을 줄이기 위해서는 도메인 일반화 기법을 활용한다. 따라서 본 논문은 위성 영상 데이터에 존재하는 도메인 간극을 줄이기 위해 RSC를 적용하여 도메인 간극의 극복 정도를 확인했다. 그 결과 표 1.에서 볼 수 있듯, F1 score의 성능이 향상된 경우가 21가지, 성능이 하락한 경우가 3가지로 성능이 향상된 경우가 압도적으로 많았다.

또한 Accuracy를 측정하여 위성 영상 데이터에서 도메인 일반화 기법의 효과를 입증했다. 그 결과 표 2.에서 볼 수 있듯, 성능이 향상된 경우가 18가지, 성능이 하락한 경우가 7가지로 성능이 향상된 경우가 더 많았다. 또한, 성능이 가장 크게 하락했을 때는 5.1%가 하락했지만, 성능이 가장 크게 올랐을 때는 31.9%가 상승한 것을 확인할 수 있다. 이를 통해 도메인 간극이 존재하는 위성 영상 데이터 집합에서 도메인 일반화 기법으로 도메인 간극을 극복한 것을 확인할 수 있다.

IV. 결론 및 논의

본 논문은 데이터 생성이 어려운 위성 영상 데이터를 자기 지도 학습에 활용하기 위해 데이터 집합 사이의 도메인 간극을 확인했으며, 나아가 도메인 일반화 기법을 통해 도메인 간극을 극복했다. 이를 통해 자기 지도 학습 기법에 도메인 일반화 기법을 함께 적용하여 학습을 진행한다면 위성 영상 학습 데이터가 적은 단점을 극복하는 것을 기대해볼 수 있다. 그렇게 된다면 많은 위성 영상 데이터에서 자기 지도 학습을 사용해 백본망(Backbone Network)을 학습한 후, 이미지 인식보다 조밀한 예측(Dense Prediction)에 해당하는 객체 탐지, 변화 탐지(Change Detection)와 이미지 분할(Image Segmentation)에서 성능 향상을 기대해볼 수 있다.

ACKNOWLEDGMENT

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

참고 문헌

- [1] kaiming H., Xiangyu Z., Shaoqing R., and Jian S., "Deep residual learning for image recognition.", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.
- [2] Huang Z., Wang H., Xing E. P., and Huang D., "Self-challenging improves cross-domain generalization.", European Conference on Computer Vision, pp. 124-140, 2020.
- [3] Kim D., Saito K., Oh T. H., Plummer B. A., Sclaroff S., and Saenko K., "Cds: Cross-domain self-supervised pre-training.", Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9123-9132, 2021.
- [4] Yang Y., and Newsam S., "Bag-of-visual-words and spatial extensions for land-use classification.", Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems, pp. 270-279, 2010.
- [5] Xia G., Hu J., Hu F., Shi B., Bai X., Zhong Y., Zhang L., Lu X., "AID: A benchmark data set for performance evaluation of aerial scene classification.", Proceedings of the IEEE Transactions on Geoscience and Remote Sensing, pp. 3965-3981, 2017.
- [6] Wang Q., Liu S., Chanussot J., and Li X., "Scene Classification With Recurrent Attention of VHR Remote Sensing Images.", in IEEE Transactions on Geoscience and Remote Sensing, vol. 57. no. 2. pp. 1155-1167, 2019.
- [7] Zhou W., Newsam S., Li C., and Shao Z., "PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval.", ISPRS journal of photogrammetry and remote sensing, vol. 145 pp.197-209, 2018.
- [8] Dai D., and Yang W., "Satellite Image Classification via Two-Layer Sparse Coding With Biased Image Representation.", in IEEE Geoscience and Remote Sensing Letters, vol. 8 no. 1 pp. 173-176, 2011.