

# 능동 학습과 Grad CAM을 활용한 고성능 분류 모델 연구

이준영, 강신재, 한동석\*  
경북대학교

ddr37105@knu.ac.kr, kangsj129@knu.ac.kr, \*dshan@knu.ac.kr

## Research on High-Performance Labeling Models Using Active Learning and Grad CAM

Jun Yeong Lee, Sin Jae Kang, Dong Seog Han\*  
KyungPook Univ.

### 요약

본 논문은 기존의 CAM (Class Activation Map) 연구들이 주로 지역화 성능에 집중했던 것과는 달리, 활성화 영역을 시각화한 후 이를 학습 정확도를 향상시키기 위한 능동 학습의 후처리 과정에 활용하는 방법을 제안한다. 저자의 이전 연구에서는 능동 학습 과정에서 새롭게 라벨링하는 데이터에 데이터 증강을 적용하면 분류 모델의 과도한 확신 문제를 완화할 수 있으며, 정확도가 낮은 클래스들을 중간 단계에서 제외하면 남은 클래스들의 학습 효율성을 크게 높일 수 있다는 결과를 도출했다. 이를 바탕으로, 본 논문은 능동 학습 과정에서 Grad-CAM의 히트맵과 마스크를 통해 얻은 활성화 점수를 새로 라벨링할 데이터 선정에 반영하는 방식을 제안한다. 해당 모델은 매우 높은 정확도를 보이는 클래스의 수를 늘리는 동시에, 학습에 필요한 라벨링 데이터의 양을 크게 줄여 추후 제품 생산 검수에 필요한 인력을 줄여 공장 자동화에도 기여할 수 있을 것으로 기대된다.

### I. 서론

최근 몇 년 동안 딥 러닝 기술의 발전은 여러 산업 분야에서 생산 자동화의 핵심 동력으로 자리 잡고 있다. 그러나 딥 러닝 모델의 복잡성과 그에 따른 블랙 박스 문제로 인해, 실제 산업 현장에서 적용되는 모델이 왜 특정한 결정을 내리는지 이해하기 어렵다는 한계가 존재한다. 이런 이유로 딥러닝 기반 자동화 시스템에서 설명 가능성의 중요성이 부각되고 있다[1].

설명 가능성을 달성하기 위한 대표적인 방법은 CAM이다. CAM은 모델이 입력 데이터의 어느 부분을 참조하여 특정 클래스를 예측하는지를 시각적으로 보여줌으로써, 모델의 의사결정 과정을 설명하는 데 도움을 준다 [2]. 이를 통해 클래스별로 학습 결과에 따른 성과를 파악하고, 부족한 부분을 보강할 수 있다.

본 논문에서는 앞선 연구를 바탕으로 CAM을 활용한 능동 학습 과정에서 더 많은 클래스의 정확도를 위해 학습 중단과 증강 복제를 통해 이미지의 특정 부분 분류에 의존도 문제를 해결하는 것을 제시한다.

### II. 본론

본 논문은 산업 현장에서 흔히 볼 수 있는 클래스 간 데이터 수의 불균형을 평가에 반영하기 위해, 특수한 COCO 데이터셋을 7:2:1 비율로 학습, 검증, 실험 세트로 나눈다. 이 과정에서 증강 표본화를 통해 데이터가 부족한 클래스들도 학습에 충분히 반영될 수 있도록 학습 모델을 구성한다. 추가로, 데이터 수가 적은 클래스가 학습에 미치는 영향을 최소화하기 위해, 모든 클래스의 데이터 수를 단순 복사하여 보완한다.

표 1은 Grad-CAM으로 측정된 점수를 클래스별로 활용해 능동 학습 과정에서 새로 라벨링할 클래스를 선별해서, 결론적으로 추가적인 후처리를 5회 적용한 학습의 클래스별 정확도를 정리한 결과이다.

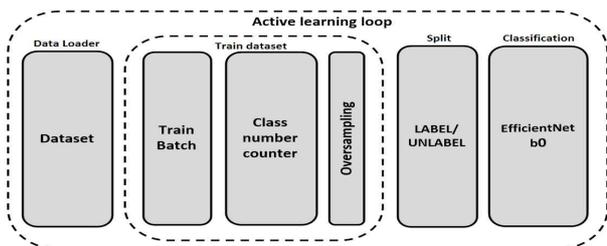


그림 1. 제시된 능동 학습의 파이프라인

표 1. 능동 학습에 Grad CAM을 통해 라벨링한 비교

Elements	Class (total = 15)	Number of labeled data	Best Accuracy (%)
Default > Proposed	1	122(+ 11*2)	95.06 > 99.99
	2	136(+ 18*2)	95.12 > 99.99
	3	212(+ 56*2)	94.86 > 99.99
	4	352(+ 126*2)	94.44 > 99.99
	5	504(+ 202*2)	92.86 > 99.99
	6	444(+ 172*2)	91.43 > 99.99
	7	264(+ 107*2)	<b>82.12 &gt; 97.56*</b>
	8	562(+ 231*2)	89.51 > 95.30
	Excluded	#	

### III. 결론

본 논문은 Grad CAM으로 계산한 활성화 점수를 능동 학습에 적용해 매우 높은 정확도를 보이는 클래스의 수를 크게 늘렸으나, 학습을 위해 라벨링할 데이터의 양은 기존에 비해서 크게 줄여 그 우수성을 증명한다.

### ACKNOWLEDGMENT

이 논문은 2023년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원(P0024162, 2023년 지역혁신클러스터육성)과 정부(과학기술정보통신부)의 재원으로 과학기술사업화진흥원의 지원을 받아 수행된 연구임('학연협력플랫폼구축 시범사업' RS-2023-00304695).

### 참고 문헌

- [1] KIM, Joongsu, et al. Deep learning based automatic defect classification in through-silicon Via process: FA: Factory automation. IEEE/29th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC). 2018. p. 35-39.
- [2] KWEON, Hyeokjun; YOON, Kuk-Jin. From SAM to CAMs: Exploring Segment Anything Model for Weakly Supervised Semantic Segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024. p. 19499-19509.