

데이터 증폭 기법과 분류 모델에 따른 배터리 결함 검출 성능 비교

김나혜¹, 정유진¹, 최문영¹, 김지열², 이상철*

대구경북과학기술원 기초학부, *대구경북과학기술원 나노기술연구부

{knh2222, eugene_jeong, cmj021205, artjy0506, sangchul.lee}@dgist.ac.kr

Performance Comparison of Battery Defect Detection According to Data Augmentation Techniques and Classification Models

Nahye Kim¹, Eugene Jeong¹, Moonyoung Choi¹, Jiyeol Kim², Sangchul Lee*

School of Undergraduate of Studies, DGIST, *Division of Nanotechnology, DGIST

요약

본 논문은 배터리 결함 검출의 정확도를 높이기 위해 다양한 데이터 증폭 기법과 분류 모델의 성능을 비교한다. 학습 데이터셋 부족 문제를 해결하기 위해 다양한 데이터 증폭 기법을 제안하고, 각 증폭 데이터셋에 대해 4 개의 대표적인 분류 모델을 적용하였다. 모델의 성능은 정확도와 AUC(Area Under the Curve) 점수로 평가되었다. 그 결과, Mix-up 증폭 기법이 결함 검출의 정확도와 AUC 점수를 유의미하게 향상시켰으며, 특히 k -Nearest Neighbors, Random Forest 에서 높은 성능을 보였다. 이를 바탕으로, 배터리 결함 검출의 정확도를 높이는 최적의 접근 방안을 제시한다.

I. 서론

리튬 이온 배터리 내부의 결함을 검출하기 위해서는 정상 데이터와 비정상 데이터를 수집해야 한다[1]. 그러나 배터리 제작에 시간이 많이 소요되기 때문에 충분한 데이터를 확보하기 어렵다. 이를 해결하기 위하여 원본 이미지에 회전, 상하-좌우 반전, 채도 조절, 이미지 합성, Mix-up 의 다섯 가지 데이터 증폭(Data Augmentation) 기법을 적용하여 데이터의 다양성을 확보하였다[2]. 본 논문은 각 증폭 기법이 최적의 분류 모델과 결합되면 어떤 결과를 보이는지 실험적으로 검증한다.

II. 배터리 자기장 이미지 데이터 증폭

2.1. 배터리 자기장 이미지 데이터 증폭

그림 1 은 본 논문에서 제안하는 데이터 증폭 기법과 모델 학습 및 평가 과정을 설명한다.

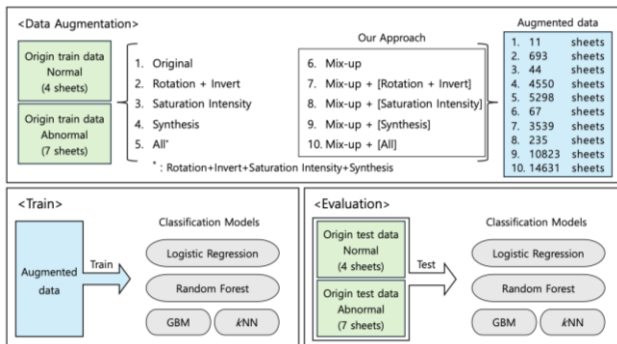


그림 1. 모델 학습 및 평가 과정

2.2. 모델 선정 및 데이터 증폭 방법

2.2.1. 모델 선정

본 논문에서는 다수의 기본적인 머신 러닝 모델을 테스트 모델로 선정하였다[3]. 이는 데이터에 대해 최적의 성능을 발휘하는 모델을 찾고, 과적합을 방지하기 위함이다.

2.2.2. 기존 데이터 증폭 방법

배터리 자기장 이미지의 채도를 조절해 전류 세기가 점진적으로 증가된 새로운 이미지를 생성한다. 상하-좌우 반전과 -10° 에서 10° 까지 1° 간격의 회전을 통해 스캔 방향의 영향을 최소화한다. 이미지 쌍을 구성하여 평균 픽셀 값을 가진 새로운 이미지를 반복적으로 합성해 데이터를 증폭한다.

2.2.3. Mix-up

Mix-up 증폭 기법을 활용하여 정상과 결함 데이터를 특정 비율로 합성한다. 정상 비율과 반대 비율을 각각 10%, 20%, 30%로 설정하여 이미지를 생성한다. 정상과 결함 데이터 각각 4 장, 7 장을 활용해 총 28 장의 이미지로 증폭을 수행한다.

III. 실험 결과

3.1. 실험 데이터

본 논문에서는 총 22 장(정상: 8, 결함: 14)의 데이터 중 11 장(정상: 4, 결함: 7)을 활용해 데이터 증폭을 수행하고, 나머지는 테스트 데이터로 사용하였다. 그림 2 는 본 논문에서 사용한 학습 데이터셋을 나타낸다.

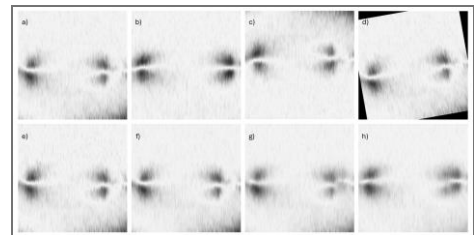


그림 2. 학습 데이터셋

¹ : these authors contributed equally to this work.

* : corresponding author

표 1. 각 분류 모델과 Mix-up 데이터셋에 따른 배터리 결함 검출 성능 평가 결과

Dataset	GBM		kNN		LR		RF		Avg	
	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC
ORIG	63.64	0.71	72.73	0.71	45.45	0.71	81.82	0.71	65.91	0.71
MIX	63.64	0.71	81.82	0.89	45.45	0.71	72.73	0.71	65.91	0.76
MIX-R+I	90.91	0.93	81.82	0.75	81.82	0.79	81.82	0.86	84.09	0.83
MIX-INT	63.64	0.71	81.82	0.82	45.45	0.71	72.73	0.71	65.91	0.74
MIX-SYN	54.55	0.71	63.64	0.36	45.45	0.71	72.73	0.71	59.09	0.62
MIX-ALL	90.91	0.93	81.82	0.93	81.82	0.75	87.88	0.89	85.61	0.88
Avg	71.22	0.78	77.28	0.74	57.57	0.73	78.29	0.77	71.09	0.76

(a). 정상 이미지, (b). 탭 결함 이미지, (c). 정상 상하 반전 이미지, (d). 정상 회전 이미지, (e). 정상 채도 조절 이미지, (f). 정상 합성 이미지, (g). 정상 70% Mix-up 이미지, (h). 탭 결함 70% Mix-up 이미지.

3.2. 데이터 증폭 방법에 따른 정확도

회전, 상하 좌우 반전, 채도 조절, 합성의 증폭 기법을 적용한 데이터셋과 Mix-up 기법을 추가로 적용한 데이터셋을 각각 분류 모델에 학습시켜 정확도와 AUC 점수를 비교하였다. 그 결과, Mix-up 기법을 적용한 데이터로 학습했을 때 정확도와 AUC 점수가 평균적으로 약 10% 향상되었다.

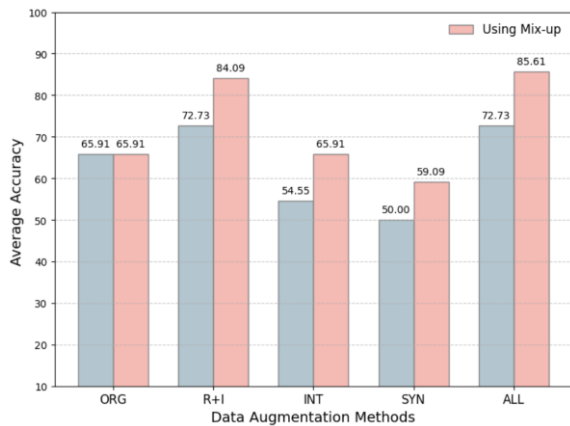


그림 3. Mix-up 기법 적용에 따른 정확도 변화

3.3. 모델 종류에 따른 정확도

4 가지 모델에 대해 데이터 증폭 기법에 따른 정확도를 평가한 결과, Mix-up 기법을 적용한 데이터로 학습했을 때 전반적으로

더 우수한 성능을 보였다. 특히, 전체 데이터 증폭 기법에 대해서는 RF 와 kNN 이 평균 정확도와 AUC 점수에서 가장 높은 값을 기록하였다.

IV. 결론

본 논문에서는 Mix-up 기법을 적용하여 자기장 이미지 데이터를 다양한 방식으로 증폭하고, 이를 학습한 분류 모델을 통해 배터리 결함 여부를 판별하는 방법을 제안하였다. Mix-up 기법을 추가로 활용한 결과, 기존 증폭 방식 대비 약 10% 향상된 성능을 나타냈으며, 여러 모델 간 비교에서는 RF 와 kNN 이 가장 우수한 성능을 보였다.

ACKNOWLEDGEMENT

이 논문은 과학기술정보통신부에서 지원하는 DGIST 기관 고유 사업(24-ET-02)과 DGIST UGRP(Undergraduate Group Research Program)(2024020031), 한국연구재단에서 지원하는 ERC 선도연구센터 사업(RS-2023-00222166)에 의해 수행되었습니다.

참 고 문 헌

- [1] Lee, M., et al. (2023). Diagnosis of Current Flow Patterns Inside Fault-Simulated Li-Ion Batteries via Non-Invasive, In Operando Magnetic Field Imaging. *Small Methods*, 7(11), 2300748.
- [2] Kim N., et al. 배터리 자기장 이미지 데이터 증폭을 통한 배터리 결함 검출. *한국정보과학회 학술발표논문집*, pp. 280-285., 2024.
- [3] Burkov, A., *The Hundred-Page Machine Learning Book*, 2019.