


1단계(3차년도) 기술문서

(과제명) 대규모 분산 에너지 저장장치 인프라의 안전한 자율운영
및 성능 평가를 위한 지능형 SW 프레임워크 개발
(과제번호) 2021-0-00077

- 결과물명 : 사전학습 딥러닝 모델을 통한 태양광 BMS 데이터의 이상감지 및 상태진단 기술 문서
- 작성일자 : 2023년 11월 15일

과학기술정보통신부 SW컴퓨팅산업원천기술개발사업
“1단계(3차년도) 기술문서” 로 제출합니다.

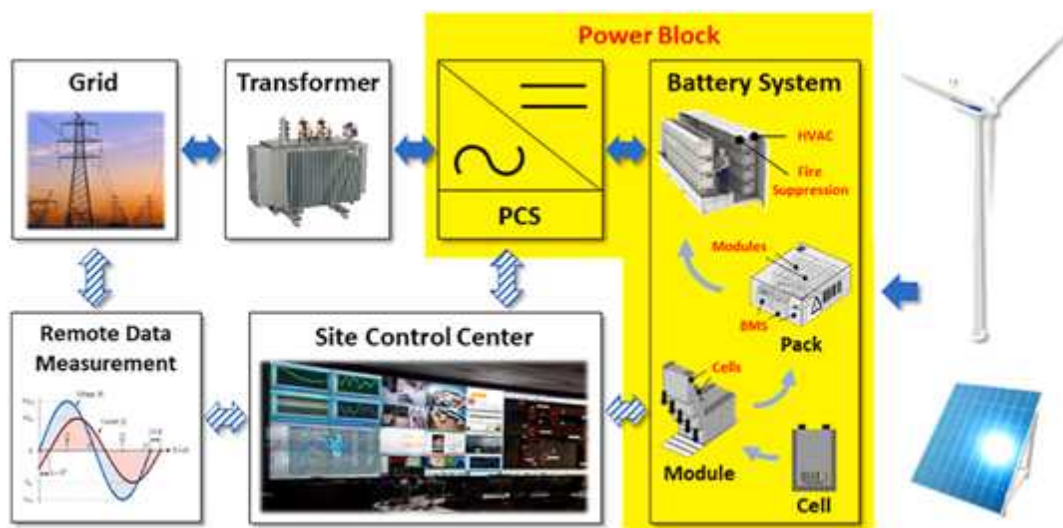
수행기관	성명/직위	확인
서울대학교	강명주/연구책임자	

정보통신기획평가원장 귀하

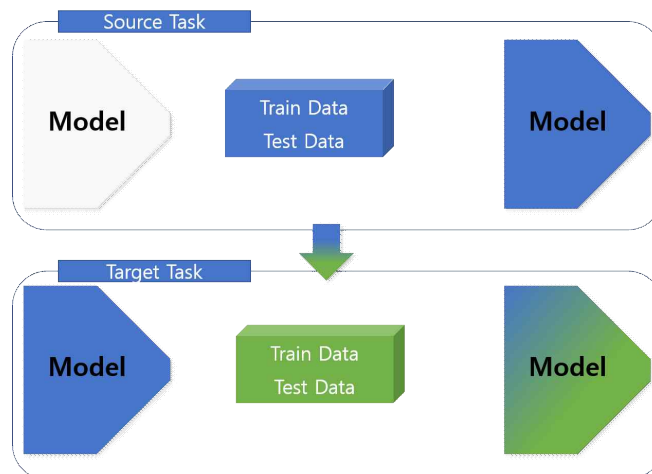
A. 사전학습 딥러닝 모델을 통한 태양광 BMS 데이터의 이상감지

1. 개요

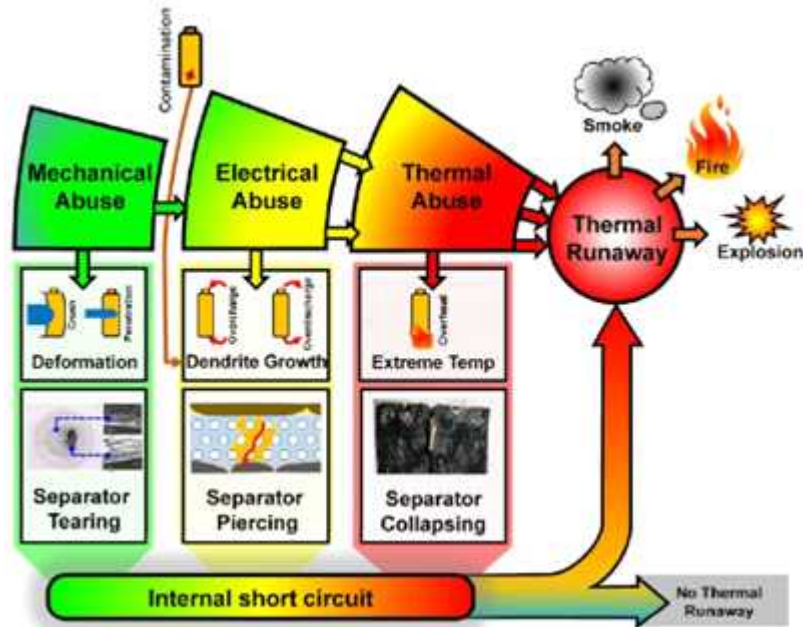
ESS(Energy storage system, 에너지 저장 시스템)는 에너지를 저장하는 데 사용되는 기술과 방법론을 총칭하며, 특히 잉여 에너지가 발생하는 발전소 등 산업 전반에서 폭넓게 활용된다. 한편 에너지 저장장치인 배터리는 기계적, 전기적, 열적 원인에 의한 열폭주(Thermal runaway)의 위험성이 있으며, 이는 화재나 폭발 등의 치명적인 사고로 이어질 수 있다. 따라서 ESS의 운영에는 이상 상황을 사전에 탐지하는 과정이 필수적이다. 본 연구에서는 딥러닝 기반 이상 탐지 모델을 이용해 여러 배터리 데이터에 적용하여 전이학습의 영향을 살펴본다. 또한 배터리가 사회 전반적으로 이용되는 데 비해 데이터의 수는 많지 않고 기업에서 가지고 있는 데이터는 공개가 쉽지 않아 다양한 연구와 개발에 한계가 많다. 이러한 데이터 부족을 타개하기 위해 딥러닝 기반 상태 진단 모델을 이용해 배터리 데이터에 적용하여 전이학습의 영향을 살펴본다.



<에너지 저장 장치 구조도>



< 전이학습의 구조 예시 >



<배터리 열 폭주의 원인과 양상>

2. 정의 및 주요 내용

○ 전이학습

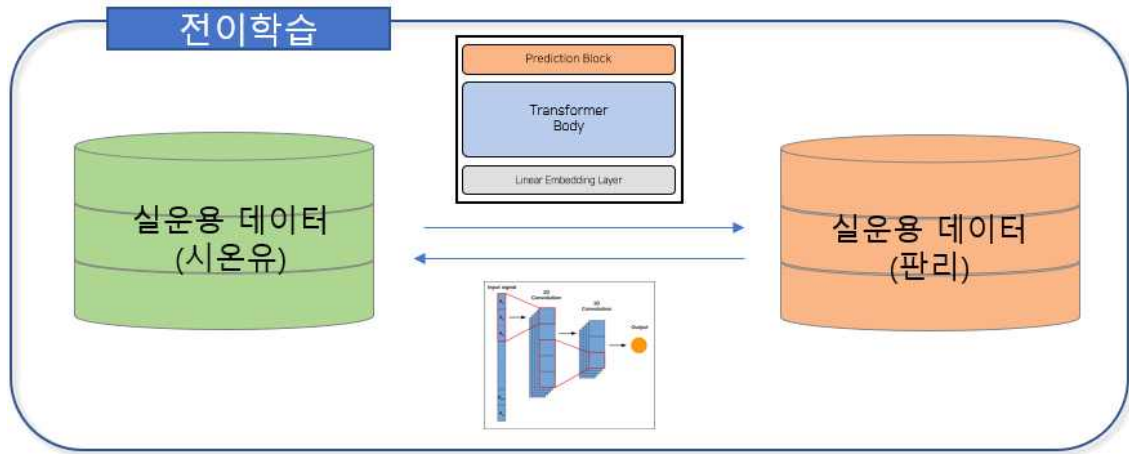
전이학습은 데이터가 부족한 상황에서 많이 쓰이며 하나의 데이터셋을 통해 학습한 모델을 또 다른 데이터를 통해 다시 한번 학습하는 것을 의미한다. 이 과정은 우리가 배웠으면 하는 특정한 데이터셋이 있지만 다른 데이터셋을 통하여 모델을 예습 혹은 선행학습을 하듯이 사전에 미리 배우고 오는 것을 의미한다. 사전 학습을 위한 데이터셋은 주로 우리가 배우길 원하는 데이터셋과 형태나 성질이 유사한 것을 주로 선정하지만 때때로 성질이 다른 데이터셋을 선정하기도 한다.

전이학습의 성공여부 및 효과를 측정하기 위해서는 크게 3가지가 사용된다. 첫 번째로는 전이학습을 했을 때의 정확도가 하지 않았을 때의 정확도 보다 높은 경우 효과가 있는 것으로 판단한다. 두 번째로는 특정 정확도까지 도달하는데 걸리는 시간이 줄어드는 경우 효과가 있다고 판단한다. 세 번째로는 두 번째와 비슷하게 특정 정확도까지 도달하는데 필요한 우리가 배우고자 하는 데이터셋의 양이 얼마나 줄어들었는지로 판단한다. 그 외에도 실험상황과 여건에 따라 연구자가 정하는 여러 기준으로 전이학습의 효과를 측정 할 수 있다.

3. 기법

이상 탐지 모델을 기반으로 실 운용 데이터들을 이용해 전이학습의 영향을 살펴볼 것이다. 먼저 한 데이터에 대해 사전학습(pre-train)을 진행하고, 그 훈련된 모델을 다른 데이터에 대해 Fine-tuning하게 된다. 전이학습의 영향을 알아보기 위해 Fine-tuning

할 때 사용한 데이터를 가지고 가중치가 초기화된 상태로 학습을 진행한다. 그래서 Fine-tuning된 모델의 학습 시간과 가중치를 초기화한 상태로 학습된 모델의 학습 시간을 비교한다. 여기서 학습 시간은 목표 정확도에 도달하는 데 걸리는 시간을 말한다.



<실제 운용 데이터를 활용한 딥러닝 전이학습 예시>

3. 실험

○ 데이터

실운동 태양광 배터리 데이터인 시온유(Sionyu)와 판리(Panli) 데이터를 사용한다. 두 데이터는 태양광 배터리 데이터이므로 하루 중 낮 동안 충전을 진행하고 해가 지면 방전을 해서 하루 중 한 번의 충전과 한 번의 방전이 기록되어 있다. 또한 일조량, 비, 구름 등의 여러 가지 기상변수들에 대한 영향이 데이터로 남아있다. 설치된 지역이 다른 두 실운동 태양광 배터리 데이터인 시온유와 판리 데이터는 지표 후보군을 만드는 실험을 하는 데에 적합한 배터리 데이터이다. 데이터에는 여러 값이 저장되어 있지만 사용하는 주된 값들은 V(전압), I(전류), T(온도), SoC(State of Charge), SoH(State of Health), ΔV (Voltage gap, 전압차)가 있다.

상태진단 전이학습은 18650 리튬 이온 배터리를 이용해 실험한 NASA 데이터셋을 사용하여 학습 데이터를 구축한다. NASA 데이터는 전압, 전류, 온도 데이터가 일정한 시간 간격으로(time-series) 기록되어있다. NASA 데이터셋은 총 28개의 데이터 파일이 있으며, 4개의 데이터 파일을 한 묶음으로 각기 다른 7가지 시나리오에 따라 실험이 진행되었다. 각각의 데이터 그룹은 일련의 충전 및 방전을 주기적으로 수행하여 배터리 건강 상태(SOH)를 체크하기 위한 기준을 제공한다.

○ 데이터 전처리

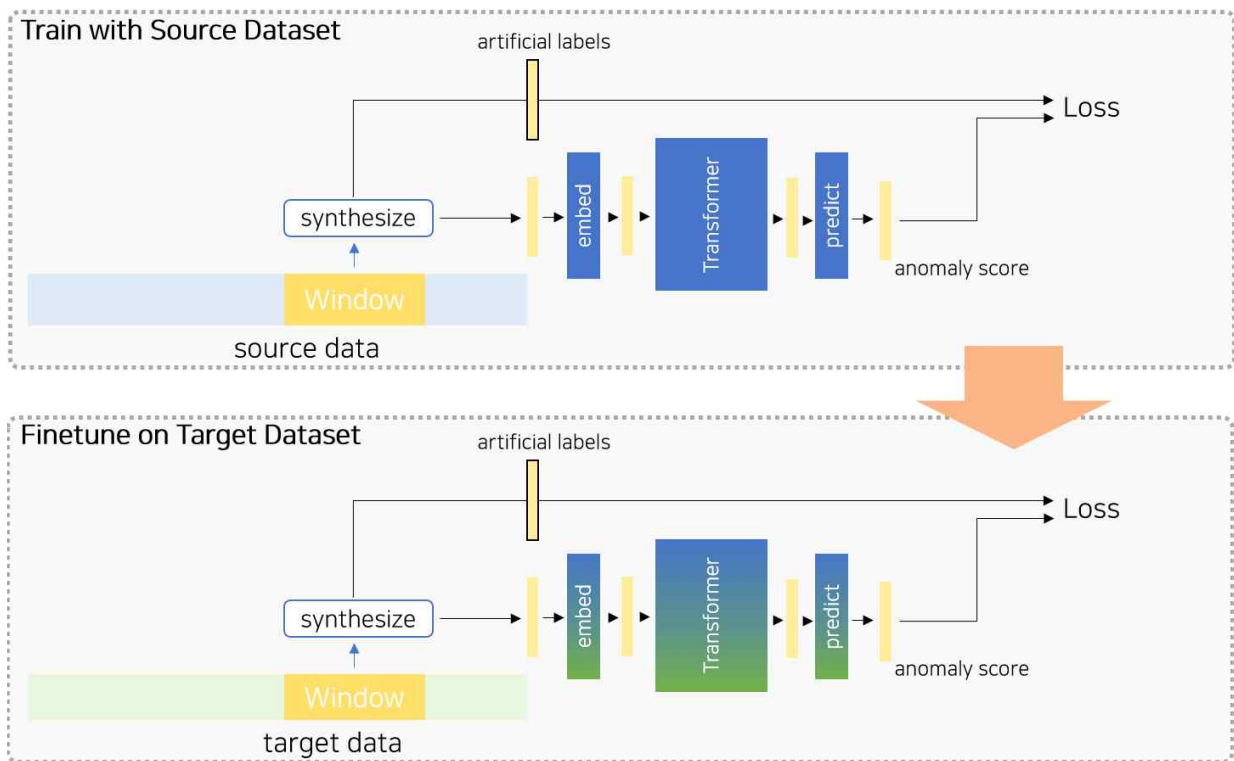
1) 정규화

데이터를 딥러닝 모델에 입력값으로 주기 위해서는 정규화 작업을 거쳐야 한다. 이 과정은 딥러닝 모델이 주어진 데이터를 더 잘 이해하고 받아들이기 위한 필수적인 과정이다. 모든 값을 정규화 작업을 진행하여 0~1 사이의 값으로 정규화 작업을 진행한다.

2) 합성 데이터

일반적으로 이상 데이터는 매우 얻기 어려우므로, 본 연구에서는 이상 탐지 모델에서 사용한 이상 데이터 합성 기법을 활용해 정상 데이터에 이상치를 합성한 합성 데이터를 사용한다.

○ 실험 방법



<이상 탐지 모델의 전이학습 실험 구성도>

-목표 : 이상 탐지 모델인 Anomaly BERT에 전이학습을 적용하였을 경우 성능 향상이 있는지 확인한다.

1. 시온유 및 관리 데이터와 합성 기법을 이용한 합성 데이터를 준비한다.
2. 가중치가 초기화된 이상 탐지 모델을 시온유 데이터를 이용해 학습한다.

3. 시온유 데이터를 이용해 학습된 모델을 Learning rate를 조절해 Fine-tuning한 다음, 판리 데이터를 이용하여 전이학습을 진행한다.
4. 목표 정확도(90%)까지 도달하는데 걸리는 학습 시간을 측정한다.
5. 가중치가 초기화된 이상 탐지 모델을 판리 데이터를 이용해 학습한다.
6. 목표 정확도(90%)까지 도달하는데 걸리는 학습 시간을 측정한 후 4번의 학습 시간과 비교하여 성능이 있음을 확인한다.

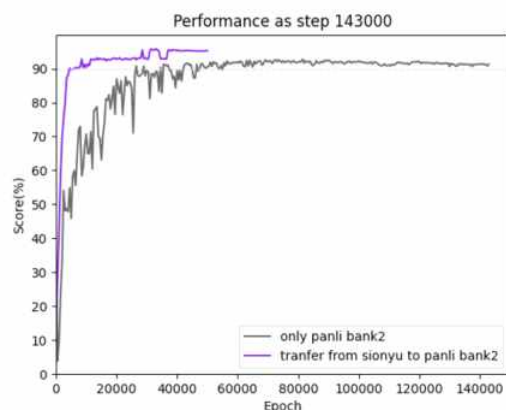
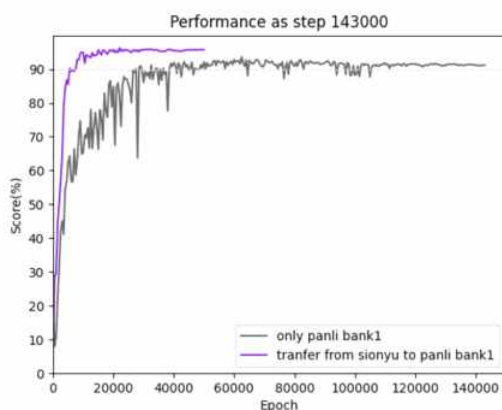
○ 실험 결과

	판리1		판리2	
	전이학습 x	전이학습 o	전이학습 x	전이학습 o
학습시간(s)	2640	528	2544	432
감소율(%)	80		83	

－ 감소율 계산 :

$$\{1 - (\text{전이학습 적용한 모델 학습시간} / \text{전이학습 적용하지 않은 모델 학습시간})\} \times 100$$

시온유 데이터로 학습 후 판리 데이터로 전이학습 하였을 시 두 데이터 모두 평균 80% 이상의 학습 시간 감소율을 보여줬다.



<이상 탐지 모델의 전이학습 정확도 비교 (좌:판리1, 우:판리2)>

4. 결론

본 연구에서는 실제 ESS 데이터인 시온유에 대해 학습을 진행한 후, 다른 ESS 데이터인 판리에 전이학습을 진행하고 전이학습의 영향을 살펴보았다. 딥러닝 기반의 이상 탐지 모델이 전이학습 시 가중치를 초기화한 상태로 학습을 진행할 때보다 학습 시간을 단축하게 한다는 사실을 입증했으며, 성능 또한 떨어지지 않는다는 것을 알 수 있다.

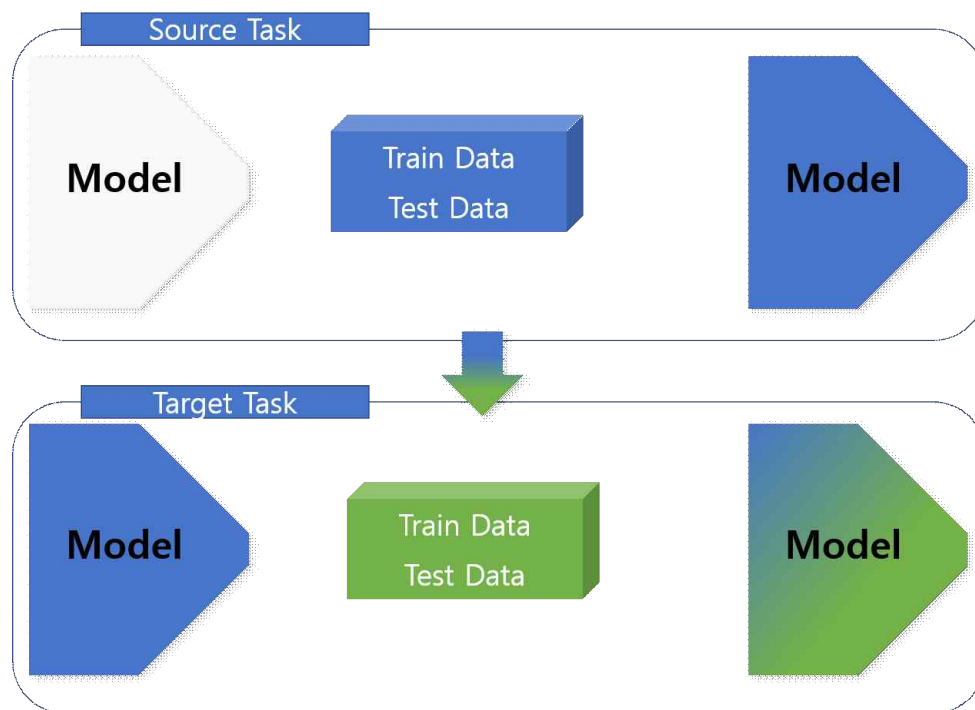
5. 참고자료

- Yungi Jeong, et al., “AnomalyBERT: Self-Supervised Transformer for Time Series Anomaly Detection using Data Degradation Scheme”, International conference on learning representations, 2023.

B. 사전학습 딥러닝 모델을 통한 태양광 BMS 데이터의 상태진단

1. 개요

ESS(Energy storage system, 에너지 저장 시스템)는 에너지를 저장하는 데 사용되는 기술과 방법론을 총칭하며, 특히 잉여 에너지가 발생하는 발전소 등 산업 전반에서 폭넓게 활용된다. 특히 배터리는 개인에서부터 기업까지 사회 전반에서 필수적이며 이러한 배터리에 대한 상태 진단은 배터리의 효율적인 사용과 안전한 운용에 필요한 가장 핵심적인 연구분야이다. 배터리가 사회 전반적으로 이용되는 데에 비해 데이터의 수는 많지 않고 기업에서 가지고 있는 데이터는 공개가 쉽지 않아 다양한 연구와 개발에 한계가 많다. 이러한 데이터 부족을 타개하기 위해 본 연구에서는 딥러닝 기반 상태 진단 모델을 이용해 배터리 데이터에 적용하여 전이학습의 영향을 살펴본다.



< 전이학습의 구조 예시 >

2. 정의 및 주요 내용

○ 전이학습

전이학습은 데이터가 부족한 상황에서 많이 쓰이며 하나의 데이터셋을 통해 학습한 모델을 또 다른 데이터를 통해 다시 한번 학습하는 것을 의미한다. 이 과정은 우리가 배웠으면 하는 특정한 데이터셋이 있지만 다른 데이터셋을 통하여 모델을 예습 혹은 선행학습을 하듯이 사전에 미리 배우고 오는 것을 의미한다. 사전 학습을 위한 데이터셋은 주로 우리가 배우길 원하는 데이터셋과 형태나 성질이 유사한 것을 주로 선정하지만 때때로 성질이 다른 데이터셋을 선정하기도 한다.

전이학습의 성공여부 및 효과를 측정하기 위해서는 크게 3가지가 사용된다. 첫 번째로는 전이학습을 했을 때의 정확도가 하지 않았을 때의 정확도 보다 높은 경우 효과가 있는 것으로 판단한다. 두 번째로는 특정 정확도까지 도달하는데 걸리는 시간이 줄어드는 경우 효과가 있다고 판단한다. 세 번째로는 두 번째와 비슷하게 특정 정확도까지 도달하는데 필요한 우리가 배우고자 하는 데이터셋의 양이 얼마나 줄어들었는지로 판단한다. 그 외에도 실험상황과 여건에 따라 연구자가 정하는 여러 기준으로 전이학습의 효과를 측정 할 수 있다.

3. 실험

○ 데이터

상태진단 전이학습은 18650 리튬 이온 배터리를 이용해 실험한 NASA 데이터셋을 사용하여 학습 데이터를 구축한다. NASA 데이터는 전압, 전류, 온도 데이터가 일정한 시간 간격으로(time-series) 기록되어있다. NASA 데이터셋은 총 28개의 데이터 파일이 있으며, 4개의 데이터 파일을 한 묶음으로 각기 다른 7가지 시나리오에 따라 실험이 진행되었다. 각각의 데이터 그룹은 일련의 충전 및 방전을 주기적으로 수행하여 배터리 건강 상태(SOH)를 체크하기 위한 기준을 제공한다.

28개의 데이터에는 셀 번호 RW1, RW2, ... , RW28(RW : Random Walk)로 이름 붙여졌다. 모든 데이터는 RW모드와 Reference 모드가 반복된다. RW 모드로 충·방전을 여러 번 반복하다가 주기적으로 Reference 모드가 수행된다. 상태진단을 위한 전이학습에는 RW17, RW18, RW19, RW20이 실험에 사용되었고 전체 RW모드의 실험조건은 다음과 같다.

셀 번호	설명	전류(A)	전압(V)
9, 10, 11, 12	실온에서의 RW 충전/방전 데이터	충/방전 -4.5 ~ 4.5	3.2 ~ 4.2
25, 26, 27, 28	40도/우편향 RW 방전 데이터	충전 2.0 방전 0.5 ~ 5.0	3.2 ~ 4.2
17, 18, 19, 20	우편향 RW 방전 데이터	충전 2.0 방전 0.5 ~ 5.0	3.2 ~ 4.2
21, 22, 23, 24	40도/좌편향 RW 방전 데이터	충전 2.0 방전 0.5 ~ 5.0	3.2 ~ 4.2
13, 14, 15, 16	좌편향 RW 방전 데이터	충전 2.0 방전 0.5 ~ 5.0	3.2 ~ 4.2
1, 2, 7, 8	충전 시간이 가변적인 RW 방전 데이터	충전 2.0 방전 0.5 ~ 4.0	3.2 ~ 4.2
3, 4, 5, 6	RW 방전 데이터	충전 2.0 방전 0.5 ~ 4.0	3.2 ~ 4.2

<표1 배터리 cell별 Random Walk 실험 조건>

Random Walk 모드는 각 시나리오에 따라 충·방전이 랜덤하게 이루어지는 것을 말한다. RW17, 18, 19, 20은 각 시나리오에 따라 4.2V까지 충전된 배터리를 랜덤하게 선택된 전류값을 주기적으로 바꿔가며 cut-off 전압인 3.2V에 도달할 때까지 방전시킨다. 여기서 전류값의 주기적 설정은 각 시나리오마다 지정되어 있는데 표에 언급된 우편향, 좌편향이 방전 전류 무작위 지정의 확률 분포를 의미한다. 우편향이면 더 높은 전류를 뽑을 확률이 높다는 것이고 좌편향이면 더 낮은 전류를 뽑을 확률이 높다는 것을 의미한다.

Reference 모드는 배터리 상태를 계측하기 위해 별도로 규칙을 지정한 것이다. 17, 18, 19, 20번 셀에선 50번의 RW 사이클마다 레퍼런스 사이클이 시행된다.

RW 모드의 구체적인 확률분포와 수치 및 레퍼런스 사이클의 규칙에 관한 설명은 NASA 측에서 제공하는 dataset 설명서 html 파일에서 확인할 수 있다.

○ 데이터 전처리

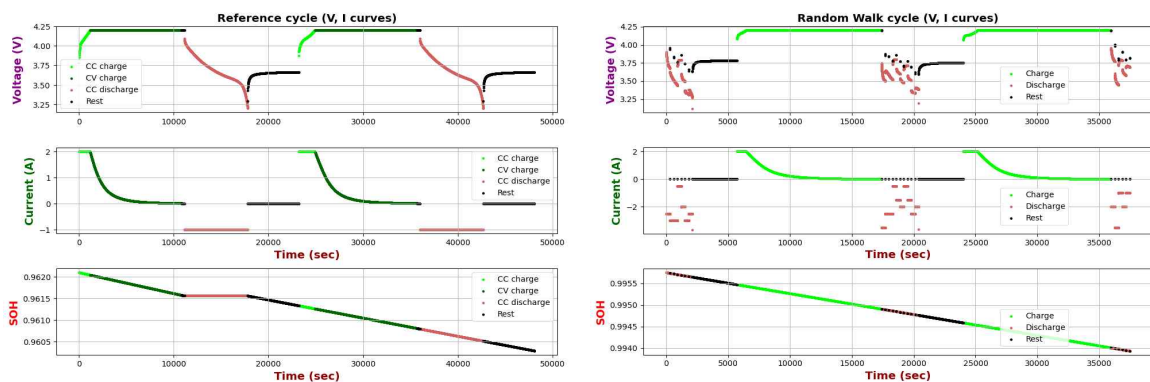
- 1) 모든 RW데이터에 포함되어있는 ‘reference discharge’ cycle을 이용하여 capacity를 계산한다. 일정한 전류값으로 방전되는 reference discharge cycle에서 capacity는 전류를 시간에 따라 적분한 값을 이용한다. 각 cycle마다 계산된 capacity가 급격히 하락하거나 상승하는 일부 비정상 구간은 정상 범위에 들 수 있게끔 평균을 내어 조정한다. reference discharge cycle이 아닌 구간에서 interpolate, extrapolate하여 각 cycle에 대한 capacity를 계산한다.

2) SOH를 계산한다. SOH는 최대 capacity값 대비 현재 capacity값의 비율로 정의한다. ($SOH = \frac{capacity}{max\ capacity}$)

3) reference cycle을 3가지 type으로 나누어 정한다. reference type A는 0.04A의 낮은 전류로 방전하는 것이다. reference type B는 CC/CV 충방전을 주어진 값에 따라 반복하는 것이다. reference type C는 주기적으로 펄스 방전을 반복하는 것이다.

4) 전하량 Q를 계산한다. 위에서 계산한 capacity에 1초 간격으로 interpolate한 전류 값을 더해 매 초마다 전하량 값을 계산하고 결측된 구간에 대해 linear interpolate 한다.

5) SOC를 계산한다. 현 시점 capacity 대비 현재 배터리의 전하량 Q의 비율로 정의한다. ($SOC = \frac{Q}{capacity}$)



<그림 : 전처리된 Reference cycle, Random walk cycle>

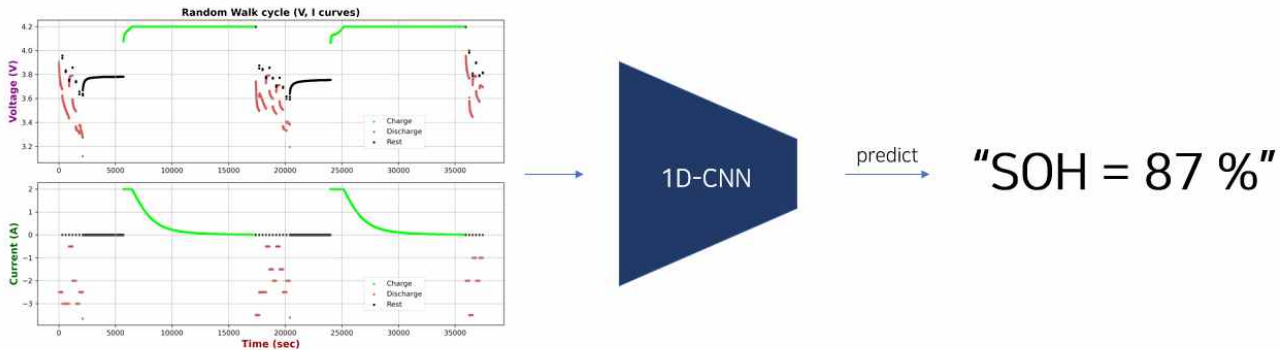
○ 실험 방법

상태진단 전이학습은 NASA RW17,18,19,20 데이터셋을 크게 2개로 나누어 실험을 구성하였다. 비교군 실험에서는 4개 데이터셋 중 2개를 이용하여 1D-CNN 모델을 통해 SOH 값을 추정한다. 실험군 실험에서는 나머지 2개 데이터셋을 이용하여 SOH 추정 학습을 진행한 후, 비교군 실험의 데이터셋에 대해 전이학습을 시행한다.

각 실험에서는 2개의 배터리 데이터 중 하나의 데이터셋을 훈련 데이터로, 나머지 하나의 데이터셋을 테스트 데이터로 사용한다.

1D-CNN 모델을 사용하여 모두 400 Epoch만큼 훈련을 시키고 Loss function으로는

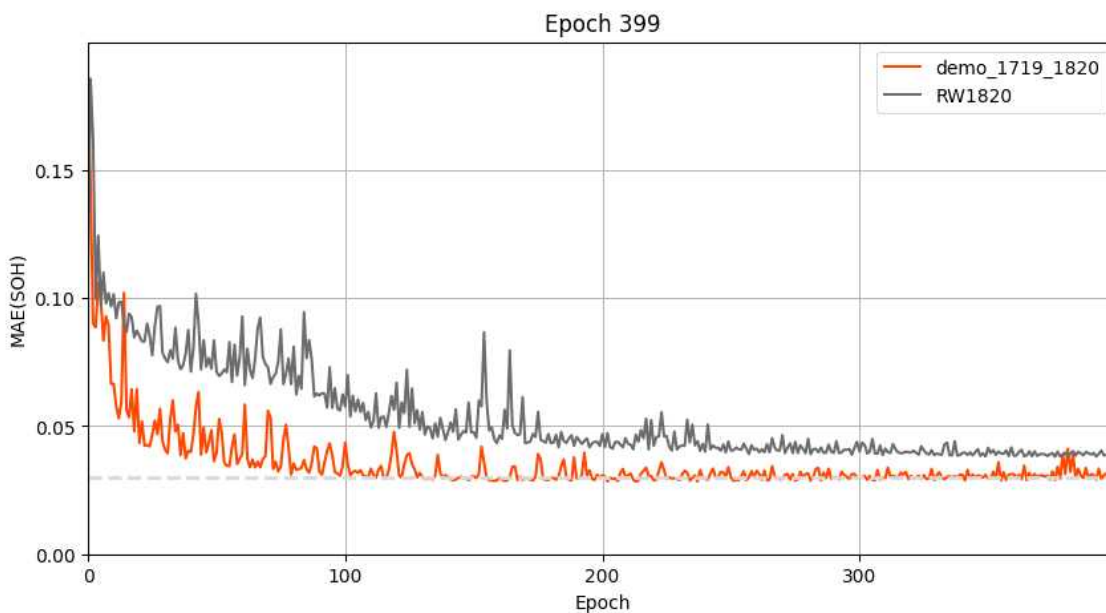
MSE (Mean Square Error) 값을 사용하고, 테스트 시에는 MAE (Mean Absolute Error)를 Evaluation metric으로 사용한다.



<그림 : SOH 추정 모델>

○ 실험 결과

전이학습을 시행한 결과와 시행하지 않은 결과를 비교한 결과 전이학습을 한 실험은 GroundTruth와 비교해 MAE값이 3%이하로 감소하였다. 실험군과 비교군에 쓰인 데이터셋을 바꿔가며 여러 차례 전이학습을 한 결과 각 학습 데이터의 에러(MAE) 감소율이 평균 20% 이상을 기록하였다.



<RW17,19 데이터에서 RW18,20데이터로 전이학습한 후 MAE>

5. 참고자료

○ Naha, Arunava, et al., "An incremental voltage difference based technique for

- online state of health estimation of li-ion batteries.”, Scientific Reports, 2020.
- Kawahara, Yohei, et al., “Development of status detection method of lithium-ion rechargeable battery for hybrid electric vehicles.”, Journal of power sources 481 (2021): 228760., 2021.
 - Venugopal, Prakash., “State-of-Health estimation of li-ion batteries in electric vehicle using IndRNN under variable load condition.”, Energies 12.22 (2019): 4338., 2019.