

Deep Learning-Based Non-Linear Prediction of Remaining Useful Life of Aircraft Engines

Min - Jung Kim[†] · Kang - Won Lee^{††}

ABSTRACT

The prediction of the remaining useful life (RUL) of aircraft turbofan engines is a critical task in prognostics and health management (PHM), as it enables the early detection of component degradation, the optimization of maintenance schedules, and the prevention of safety incidents. Recent deep learning-based RUL prediction studies have made significant progress. However, most efforts have focused on improving model architectures, while relatively little attention has been paid to the design of RUL labeling functions. This study proposes a novel approach that preserves the existing pre-training structure used in recent research while replacing the target RUL labeling function with a non-linear concave function that more accurately reflects actual degradation patterns. Using the NASA C-MAPSS (Commercial Modular Aero-Propulsion System Simulation) dataset, we evaluate a CAE (convolutional autoencoder)-RNN-based prediction model under various parameter settings and demonstrate that the proposed non-linear labeling model improves prediction accuracy in terms of RMSE (root mean squared error) and simultaneously enhances the NASA S-score safety metric compared to the conventional piecewise-linear model.

Keywords : Remaining Useful Life Prediction, Prognostics and Health Management (PHM), Deep Learning, Aircraft Turbofan Engines, C-MAPSS dataset

딥러닝 기반 비선형 모델을 이용한 항공기 엔진의 수명 예측

김민정[†] · 이강원^{††}

요약

항공기 터보팬 엔진의 잔여 유효 수명(remaining useful life, RUL) 예측은 부품 열화로 인한 고장을 사전에 탐지하여 정비 일정을 최적화하고 안전사고를 예방하는 예지보전(prognostics and health management, PHM)의 핵심 과제이다. 최근 딥러닝 기반의 RUL 예측 연구가 활발히 진행되었으나, 대부분의 연구는 모델 구조 향상에 집중되어 있으며, RUL 라벨 설계에 관심을 두지 않았다. 본 연구는 엔진 열화의 비선형적 열역학 특성을 반영하여 최신 연구에서 사용하는 사전학습 구조는 그대로 유지하되, 학습 목표인 RUL 라벨 함수만을 실제 열화 패턴과 유사한 비선형 오목(concave) 함수로 대체하는 새로운 접근법을 제안한다. NASA C-MAPSS 데이터셋을 사용하여 CAE(convolutional autoencoder)-RNN 구조에 기반을 둔 예측 모델을 다양한 파라미터로 평가한 결과, 제안한 비선형 라벨링 모델이 기존 조각선형(piecewise-linear) 모델에 비해 RMSE(root mean squared error) 기준 예측 정확도와 NASA S-score 안전성 지표를 모두 향상시키는 것을 확인하였다.

키워드 : 수명 예측, 예지보전, 딥러닝, 항공기 터보팬 엔진, C-MAPSS 데이터셋

1. 서론

항공기 터보팬 엔진의 잔여 유효 수명(remaining useful life, RUL) 예측은 부품 열화로 인한 고장을 사전에 탐지하여 정비 일정을 최적화하고, 예기치 못한 정비 중단 및 안전사고를

예방하기 위한 예지보전(prognostics and health management, PHM)의 핵심 과제이다. PHM은 설비 상태를 예측함으로써 다운타임(downtime)을 최소화하고 유지보수 계획을 최적화하는 것을 목표로 하며, 그 가운데 RUL 예측이 전체 PHM에서 중요한 역할을 담당한다. RUL 예측 연구를 위해, NASA에서 제공하는 C-MAPSS(Commercial Modular Aero-Propulsion System Simulation) 데이터셋이 표준 벤치마크로 널리 활용되어 왔고[1, 2], 이 데이터셋을 대상으로 CNN, LSTM, GRU 등 다양한 모델이 활발히 적용되면서 예측 성능이 꾸준히 향상되어 왔다[3, 4]. 그러나 지금까지 RUL 라벨 설계 방법은 대

※ 본 연구는 2025년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음(2025-0-00037)

[†] 준회원 : 세종대학교 컴퓨터공학과 학사

^{††} 종신회원 : 세종대학교 컴퓨터공학과 교수

Manuscript Received : September 30, 2025

Accepted : October 10, 2025

* Corresponding Author : Kang - Won Lee(kangwon.lee@sejong.ac.kr)

부분 초기 안정 구간을 고정 상한으로 두고 이후를 조각선형(piecewise-linear)으로 감소시키는 방식을 사용해 왔다[5, 6].

본 연구는 이러한 기존의 라벨링 방식이 실제 물리적 열화 과정과의 불일치를 가져올 수 있으며, 학습 과정에서 편향을 유발하여 예측 성능을 저해할 수 있다는 점에 착안하였다. 실제 물리 시스템의 열화는 마모, 균열, 피로 누적 등의 복합적인 요인에 의해 초기에는 완만하게 진행되더라도 수명 말기에 점점 가속화되는 비선형적(non-linear) 패턴을 따르는 경우가 일반적이다. 따라서, 단순화된 조각선형 라벨은 고장 직전의 위험 구간에서 RUL을 늦게 경보하는 경향을 보일 수 있으며, 이는 조기 경보 실패에 더 큰 패널티를 부과하는 안전 중심의 평가지표 NASA S-Score에서 높은 패널티를 받을 수 있다.

이러한 문제를 해결하기 위해, 본 연구는 기존 연구의 사전 학습 구조는 그대로 유지하면서, 학습 목표인 RUL 라벨 함수만을 실제 열화 패턴을 더 잘 모델링할 수 있는 비선형 오목(concave) 함수로 대체하는 새로운 라벨링 방식을 제안한다. 라벨의 곡률을 제어하는 핵심 하이퍼파라미터 k 는 사전 정의된 후보 집합 내에서 교차검증을 통해 데이터 기반으로 최적의 값을 자동으로 선택하도록 설계하였다.

제안 방법론의 성능은 C-MAPSS FD001 데이터셋을 사용하여 검증하였다. 성능 평가는 RMSE(root mean squared error), MAE(mean absolute error), NASA S-Score 지표를 통해 기존 조각선형 라벨 방식과의 비교 실험으로 수행하였으며, 제안한 비선형 오목 라벨 방식이 기존 방식 대비 예측 정확도와 조기 경보 성능 모두에서 우수함을 확인하였다.

본 연구의 기여는 다음과 같다:

(1) 기존 연구가 복잡한 딥러닝 모델 아키텍처를 통해 성능을 향상시키고자 한 것과 달리, 본 연구는 RUL 라벨과 실제 물리적 열화 패턴 간의 불일치가 근본적인 성능 저하 요인임을 밝힌다. 이를 바탕으로 문제의 초점을 모델 구조 최적화가 아닌 라벨 함수 설계로 전환하는 연구 방향을 제시한다.

(2) 실험을 통해 본 연구에서 제안한 모델의 성능이 기존 연구에서 보고된 성능보다 우수함을 보인다. FD001 데이터셋에서 기존의 Baseline 라벨링 방식은 GRU 모델과 결합되었을 때 16.459의 RMSE를 기록한 반면, 본 연구에서 제안한 Exp-Concave 라벨링은 14.515의 RMSE를 달성하여 성능 향상을 보였다.

2. 기존 연구

최근 예지보전 분야에서는 인공지능을 활용하여 터보팬 엔진과 같은 복잡한 산업 설비의 잔여수명을 정확하게 예측하려는 연구가 활발히 진행되고 있다. 기존 연구는 크게 물리 기반(physics-based), 데이터 기반(data-driven), 그리고 하이브리드(hybrid) 접근법으로 구분된다.

드(hybrid) 접근법으로 구분된다.

물리 기반 방법은 시스템의 동작 원리를 수학적으로 모델링하여 열화 과정을 설명하지만, 실제 환경의 복잡성과 불확실성 모델링의 한계로 인한 제약이 있다[7]. 데이터 기반 방법은 센서 데이터로부터 통계적 또는 기계학습 기법을 활용하여 열화 패턴을 학습하며, 물리 모델링을 직접적으로 하지 않고 데이터만으로 예측하기 때문에 활용 분야가 넓다[8]. 하이브리드 방법은 두 접근법의 장점을 결합하여 예측 정확도와 일반화 능력을 동시에 향상시키는 것을 목표로 한다[9].

초기 연구에서는 CNN(Convolutional Neural Network)을 활용하여 다변량 센서 시계열 데이터로부터 공간적(spatial) 특성을 추출하고 이를 RUL 예측에 적용하는 방법이 제안되었다. Li 등은[10] NASA C-MAPSS 벤치마크 데이터를 활용한 CNN 기반 RUL 회귀 모델을 통해 RMSE 약 18~20 사이클 수준의 성능을 달성하였다. LSTM(Long Short-Term Memory) 기반의 시계열 모델은 열화 궤적(degradation trajectory)의 장기 의존성(long-term dependency)을 효과적으로 학습할 수 있어 활발히 연구되었으며[11-13] CNN 기반 방법에 비해 RMSE 기준으로 더 우수한 성능을 보였다. 그러나 LSTM은 긴 시계열 처리 과정에서 그래디언트 소실(vanishing gradient) 문제가 발생할 수 있고, 실시간 적용 시 계산량이 크다는 한계가 있다.

Attention-LSTM 모델은 입력 시계열 중 중요한 시간 구간에 가중치를 부여하여 예측 정확도를 높이며, 기존 LSTM 기반 모델보다 우수한 성능을 보였다. 최근 CNN-LSTM-Attention[14], CNN-BGRU-Self-Attention[15] 등 다양한 하이브리드 모델이 제안되었다.

Elsherif 등은[16] Convolutional Autoencoder(CAE)로 시계열 데이터의 특징을 추출하고 이를 Attention 메커니즘이 결합된 LSTM에 전달하는 CAELSTM 하이브리드 모델을 제안했다. 이들은 CAE를 통해 노이즈를 줄이고 핵심적인 시간적 패턴을 효율적으로 학습시킨 후, Attention을 통해 RUL 예측에 중요한 시퀀스에 집중함으로써 FD001 데이터셋에서 낮은 RMSE를 달성하며 우수한 성능을 보였다.

이처럼 정교한 모델 구조를 통해 예측 정확도를 높이는 접근은 RUL 예측 연구의 중요한 흐름을 형성했다. 하지만 이러한 연구들 역시 학습 목표가 되는 RUL 라벨은 기존의 조각선형(piecewise-linear) 모델을 사용함으로써, 실제 물리적 열화 패턴을 정확하게 반영하지 못한다. 본 연구는 이러한 라벨 설계의 문제점이 RUL 예측 모델의 성능을 제약하는 중요한 요인 중 하나라고 판단하고, 이를 개선하기 위한 새로운 라벨링 기법을 제안한다.

3. 제안 방법

3.1 C-MAPSS 데이터셋

C-MAPSS(Commercial Modular Aero-Propulsion System

Table 1. C-MAPSS Dataset Summary

Datasets	FD001	FD002	FD003	FD004
No. of engines for training	100	260	100	249
No. of engines for testing	100	259	100	248
Operation conditions	1	6	1	6
Fault mode number	1	1	2	2
Training sample	20,631	53,759	24,720	61,249
Test sample	13,096	33,991	16,596	41,214

Table 2. Operational Parameters

Operational parameter	Description
Tr	Throttle resolver angle (TRA)
Alt	Altitude
MN	Mach number

Table 3. Simulator Inputs: Fuel Flow and 13 Health Parameters

Name	Symbol
Fuel flow	W_f
Fan efficiency modifier	fan_eff_mod
Fan flow modifier	fan_flow_mod
Fan pressure-ratio modifier	fan_PR_mod
LPC efficiency modifier	LPC_eff_mod
LPC flow modifier	LPC_flow_mod
LPC pressure-ratio modifier	LPC_PR_mod
HPC efficiency modifier	HPC_eff_mod
HPC flow modifier	HPC_flow_mod
HPC pressure-ratio modifier	HPC_PR_mod
HPT efficiency modifier	HPT_eff_mod
HPT flow modifier	HPT_flow_mod
LPT efficiency modifier	LPT_eff_mod
LPT flow modifier	LPT_flow_mod

Simulation)는 MATLAB/Simulink 기반의 엔진-제어 폐루프 (closed-loop) 시뮬레이터로, 실제 항공기 터보팬 엔진의 동역학 및 열화를 모사한다. 이 시뮬레이터는 팬(Fan), 저압/고압 압축기(LPC/HPC), 고압/저압 터빈(HPT/LPT) 등 주요 구성 요소에 대해 효율 · 유량 보정계수를 시간에 따라 열화시키며 고장 임계치까지의 다변량(multivariate) 센서 시계열 데이터를 생성한다.

C-MAPSS 데이터는 Table 1에 보이는 것과 같이 고장 모드 수(1 또는 2)와 운영 조건 수(1 또는 6) 조합에 따라 FD001~FD004 네 가지 데이터 서브셋으로 구성된다. 모든 서브셋은 각 유닛의 전 수명(run-to-failure)을 포함하는 훈련 데이터와 고장 이전 임의 시점에서 절단된 시계열로 구성된 테스트 데

Table 4. Simulator Outputs: 21 Sensor Measurements

Sensor number	Symbol	Description	Unit of measure
S1	T2	Total temperature at fan inlet	$^{\circ}\text{R}$
S2	T24	Total temperature at LPC outlet	$^{\circ}\text{R}$
S3	T30	Total temperature at HPC outlet	$^{\circ}\text{R}$
S4	T50	Total temperature at LPT outlet	$^{\circ}\text{R}$
S5	P2	Pressure at fan inlet	psia
S6	P15	Total pressure in bypass-duct	psia
S7	P30	Total pressure at HPC outlet	psia
S8	Nf	Physical fan speed	rpm
S9	Nc	Physical core speed	rpm
S10	epr	Engine pressure ratio (P50/P2)	-
S11	Ps30	Static pressure at HPC outlet	psia
S12	phi	Ratio of fuel flow to Ps30	pps/psi
S13	NRf	Corrected fan speed	rpm
S14	NRc	Corrected core speed	rpm
S15	BPR	Bypass Ratio	-
S16	farB	Burner fuel-air ratio	-
S17	htBleed	Bleed Enthalpy	-
S18	Nf_dmd	Demanded fan speed	rpm
S19	PCNfR_dmd	Demanded corrected fan speed	rpm
S20	W31 HPT	HPT coolant bleed	lbm/s
S21	W32 LPT	LPT coolant bleed	lbm/s

이터로 구분된다. 시뮬레이터 입력에는 운항 조건 3개(Table 2)와 연료 유량 및 5개의 회전체 모듈(Fan, LPC, HPC, HPT, LPT)에 대한 헬스 파라미터 13종(Table 3)을 사용한다.

시뮬레이션 모델은 Table 4와 같이 여러 센서 응답과 운용 여유도(margins)를 출력하며, 이 중 21개 센서값이 공개 데이터로 사용된다.

3.2 예측 네트워크 모델링

본 연구에서는 Elsherif 등[16]의 CAE-RNN 계열을 기반으로 예측 네트워크를 구성한다. 이 모델은 다변수 시계열 데이터로부터 RUL을 효과적으로 예측하기 위해, (1) Convolutional Autoencoder(CAE)를 이용한 특징 추출(feature extraction), (2) 순환 신경망(RNN; GRU, BiLSTM) 기반의 예측기, (3) 완전 연결 계층(Fully-Connected Layer)을 통한 최종 RUL 회귀 단계로 구성되는 3단계 구조를 갖는다.

먼저, CAE는 다변수 시계열 센서 신호로부터 노이즈를 억제하고, 엔진 열화와 관련된 중요한 시간적, 공간적 패턴을 저차원의 잠재 시퀀스 h 로 추출한다. 사전학습이 완료된 인코더 가중치는 고정되어 공통 특징 추출기로 사용된다. 이후, 추출된 잠재 시퀀스 h 는 RNN 기반 예측기로 전달되어 시계열의 장기 의존성(long-term dependency)을 학습한다. 마지막으로, 예측기의 출력은 완전 연결 계층을 거쳐 최종적으로 RUL 값을 회귀 방식으로 예측한다. 모델의 전체 흐름은 (1) CAE 사전학습 → (2) RNN 기반 예측기 → (3) FC 회귀로 구성된다.

1) Convolutional Autoencoder (CAE, 사전학습)

다변수 센서 시계열 데이터로부터 잠재 시퀀스 h 를 학습하기 위해 1D Convolutional Autoencoder(CAE)를 사용한다. 입력 윈도우 X 는 인코더 f_{enc} 를 통과하여 잠재 시퀀스 h 로 압축되며,

$$h = f_{enc}(X) = Conv(X) \quad (1)$$

디코더 f_{dec} 는 h 로부터 재구성된 \hat{X} 를 생성하고, \hat{X} 가 원본 입력 X 를 잘 복원하도록 $MSE(\hat{X}, X)$ 손실을 최소화한다. 디코더는 사전학습에만 사용되며, 예측 단계에서는 인코더의 출력 h 만이 RNN 예측기로 전달된다. 인코더/디코더의 구체 설정은 Table 5, 6에 정리하였다.

$$\hat{X} = f_{dec}(h) = Deconv(h) \quad (2)$$

Table 5. Autoencoder Encoder Layers

Layers	Parameters
Conv1D	Filters=128, kernel=3, Activation=ReLU, Padding=Same
Conv1D	Filters=128, kernel=3, Activation=ReLU, Padding=Same
Conv1D	Filters=256, kernel=3, Activation=ReLU, Padding=Same
MaxPooling1D	Pool Size=2, Padding=Same

Table 6. Autoencoder Decoder Layers

Layers	Parameters
Conv1D	Filters=128, kernel=3, Activation=ReLU, Padding=Same
Conv1D	Filters=128, kernel=3, Activation=ReLU, Padding=Same
Conv1D	Filters=128, kernel=3, Activation=ReLU, Padding=Same
UpSampling1D	Size=2

2) RNN 기반 예측기

사전 학습된 CAE(Convolutional Autoencoder)로부터 추출된 잠재 시퀀스 h 의 시간적 의존성을 모델링하기 위해 RNN(Recurrent Neural Network) 기반의 예측 모델을 사용하였다. 본 연구에서는 단방향(unidirectional) 구조의 GRU(Gated Recurrent Unit)와 양방향(bidirectional) 구조의 BiLSTM(Bidirectional Long Short-Term Memory) 아키텍처를 적용하여 예측 성능을 비교 평가하였다.

3) 완전 연결 회귀층

RNN 계층에서 생성된 최종 문맥 벡터(context vector)는 RUL 값을 예측하기 위한 완전 연결 회귀층으로 입력된다. 이 회귀층은 두 개의 ReLU 활성화 함수를 갖는 은닉층과 하나의 선형(linear) 출력층으로 구성되며, 변환된 라벨 공간에서 회귀를 수행한 뒤 역변환을 통해 원단위 RUL로 복원하여 평가한다.

3.3 데이터 전처리

1) 센서 선택(Sensor Selection)

C-MAPSS 데이터셋은 21개의 센서 측정값과 3개의 운용 설정(operational settings)값을 포함한다. 이 중 엔진 번호나 운용 사이클과 같은 식별 정보와 본 연구의 초점인 열화 상태와 직접적 관련성이 낮은 운용 설정값은 모델의 입력 특징에서 제외하였다. 21개의 센서 데이터 중에서도, 일부 센서는 Fig. 1과 같이 운용 시간 전반에 걸쳐 값이 변하지 않거나 거의 일정한 값을 유지하여 RUL 예측에 유의미한 정보를 제공하지 않는다. 이러한 불필요한 특징(feature)은 모델 학습을 방해할 뿐만 아니라, 후속 정규화 과정에서 표준편차가 0에 근사하여 분모가 0이 되는 수치적 불안정성을 야기할 수 있다.

따라서, Fig. 2와 같이 열화 경향과 유의미한 상관관계를 보이는 14개의 센서(2, 3, 4, 7, 8, 9, 11, 12, 13, 14, 15, 17, 20, 21)만을 최종적으로 선택하여 모델의 입력 특징으로 사용하였다.

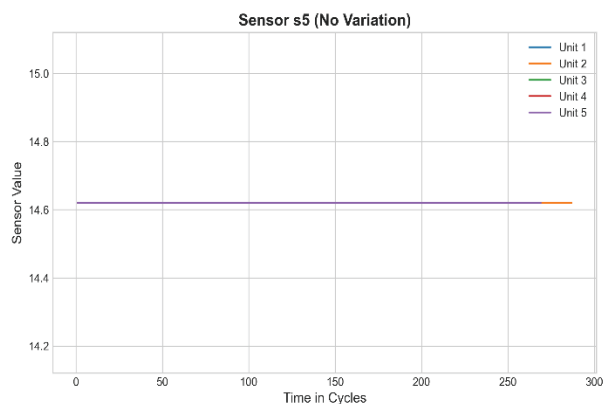


Fig. 1. Sensor 5: No Meaningful Change

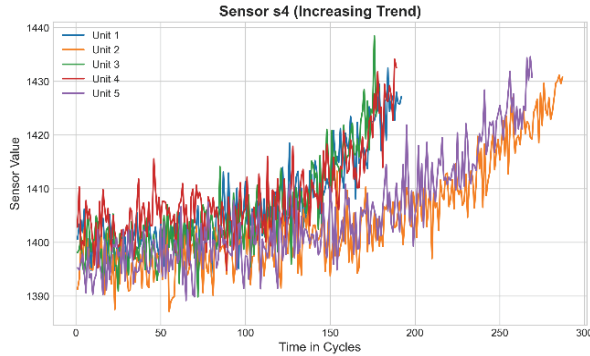


Fig. 2. Sensor 4: Increasing Trend

2) 데이터 정규화(Data Normalization)

선택된 14개의 센서는 각각 다른 물리적 단위(온도, 압력, 속도 등)와 값의 범위를 가지므로, 이를 표준화하여 모델이 모든 특징을 동등한 중요도로 학습하고 안정적으로 수렴하도록 할 필요가 있다. 본 연구에서는 각 센서 특징이 평균 0, 표준편차 1을 갖도록 변환하는 Standard Scaler(Z-score 정규화)를 식 (3)과 같이 사용하였다.

$$Z = \frac{x - \mu}{\sigma} \quad (3)$$

여기서 Z는 정규화된 값, x는 센서의 원본 측정값, μ 는 해당 센서 특징의 평균값, σ 은 해당 센서 특징의 표준편차를 나타낸다.

3.4 RUL Function

현실의 열화 곡률을 반영한 비선형 라벨이 예측 성능을 높인다는 가설을 검증하기 위해, 동일한 모델 구조와 학습·평가 절차에서 라벨링 함수만 교체하여 성능을 비교 분석하였다. 초기 탐색 단계에서는 실제 물리적 열화가 보이는 곡률을 모사하기 위해 (i) 지수형(exponential), (ii) 멱함수(power), (iii) 여러 개의 제어점(knot)을 가지며 각 구간의 기울기가 다른 조각선형(piecewise-linear)을 고려하였다. 그 결과 지수 및 멱함수와 같은 볼록(convex) 형태의 라벨은 장시간 지속되는 초기 완만 구간에 모델을 과적합시키는 경향을 보였다. 이로 인해 고장 직전의 급격한 열화 상태 변화에 대한 예측이 지연되었고, 이는 결과적으로 NASA S-score에서 큰 패널티를 받는 원인이 되었다. 반면, 조각선형 라벨은 Optuna를 사용해 제어점을 최적화할 경우 높은 표현력을 가질 수 있었으나, 탐색할 파라미터 수가 많아 최적화 과정이 불안정하고 실험 전반에 걸쳐 일관된 성능 향상을 보장하기 어려웠다.

이러한 문제들을 바탕으로, 오목(concave) 형태의 함수를 대안으로 검토하였다. 그중에서도 실제 열화 패턴, 즉 초기 완만 단계에서 말기 급격한 악화로 이어지는 곡선과 형태가 유사하면서, 단일 하이퍼파라미터 k만으로 곡률을 제어할 수 있는 지

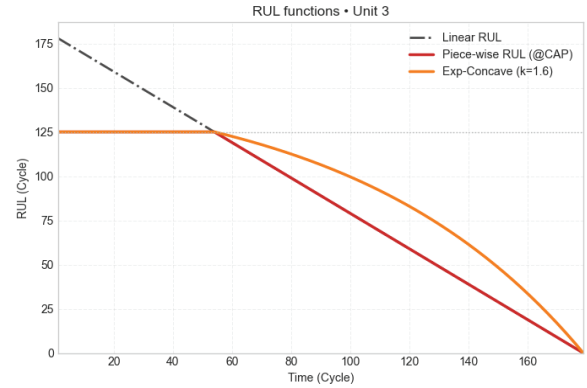


Fig. 3. Exponential-Concave functions

수-오목(Exponential-Concave) 함수를 최종 RUL 함수로 채택하였다. 대표적인 예시로 Unit 3 엔진에 대한 지수-오목 라벨 곡선을 선형 및 조각선형 라벨과 함께 Fig. 3에 나타내었다.

본 연구에서 제안하는 RUL 함수의 입력을 정의하기 위해, 원본 RUL(RUL_{raw})을 0과 1 사이의 값으로 정규화한 잔여수명 비율 u 를 도입한다. 이때 $u=0$ 은 고장 시점을, $u=1$ 은 최대 수명 상태(정상)를 의미하며, u 는 식 (4)와 같이 원본 RUL 값을 상한값(CAP)으로 먼저 제한하고, 그 결과를 다시 CAP으로 나누어 계산한다.

$$u = \frac{\text{dip}(RUL_{raw}, 0, CAP)}{CAP} \quad (4)$$

이렇게 정규화된 u 를 최종 RUL 값으로 변환하기 위해, 식 (5)와 같은 지수-오목 함수를 제안한다.

$$f_{RUL}(u; k) = CAP \cdot \frac{1 - e^{-uk}}{1 - e^{-k}} \quad (5)$$

이 함수는 경계조건 $f_{RUL}(0; k)=0$, $f_{RUL}(1; k)=CAP$ 을 정확히 만족하며, u 가 감소함에 따라 초반에는 완만하고 말기에 급격히 하락하는 오목 곡선을 형성한다. 곡선의 형태를 결정하는 하이퍼파라미터 k 는 Optuna(단일 파라미터 탐색)로 추정하였다. 실험 결과, 제안한 방식은 baseline(piecewise-linear) 방식 대비 지연 예측(late prediction)을 효과적으로 감소시켰으며, RMSE와 NASA-Score 지표 모두에서 일관되게 동등하거나 우수한 성능을 달성함을 확인하였다.

4. 평가 결과

본 연구에서 제안하는 비선형 RUL 라벨링 프레임워크의 유효성을 검증하기 위해, NASA C-MAPSS FD001 데이터셋을 대상으로 두 가지 RNN 계열 모델(BiLSTM, GRU)을 사용하여 종합적인 성능 평가를 수행하였다. 각 모델에 대해 Baseline

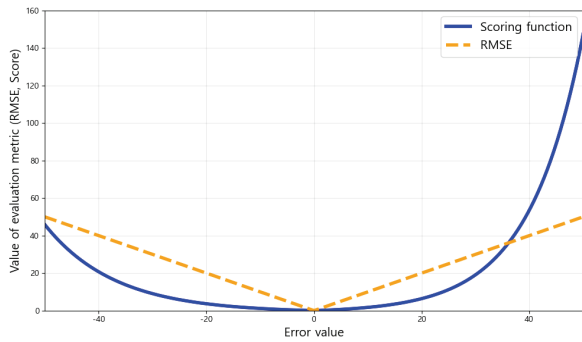


Fig. 4. RMSE versus Scoring function

(조각선형, CAP=125)과 제안된 Exp-Concave(k=1.6) 라벨을 동일한 학습·평가 프로토콜로 적용하고, RMSE와 MAE, NASA S-Score 지표를 비교하였다(Table 7-9).

NASA S-Score는 예측 오차에 비대칭 패널티를 부과하는 안정성 지표로, 실제값보다 늦게 예측하는 경우에 더 큰 패널티를 부여한다. 계산식은 (6)과 같다.

$$Score = \begin{cases} \sum_{i=1}^n (e^{-\left(\frac{d_i}{13}\right)} - 1), & \text{for } d_i < 0 \\ \sum_{i=1}^n (e^{\left(\frac{d_i}{10}\right)} - 1), & \text{for } d_i \geq 0 \end{cases} \quad (6)$$

여기서 n 은 테스트 유닛의 개수, d_i 는 i 번째 유닛의 예측 오차($d_i = \text{예측 RUL} - \text{실제 RUL}$)를 나타낸다.

Fig. 4는 S-Score와 RMSE의 오차값에 따른 패널티 변화를 비교한 것으로, S-Score가 지연 예측에 대해 RMSE보다 훨씬 큰 패널티를 부과함을 보여준다.

평가 방식은 두 가지로 구성하였다. 첫째, RUL이 125를 초과하는 구간을 125로 절단(clipping)하여 학습·검증·테스트 전 구간의 스케일을 통일한 @CAP 평가 방식, 둘째, 원본(raw) RUL 라벨을 그대로 사용한 raw 평가 방식이다. 본 연구에서는 이 두 가지 평가 기준을 통해 모델의 성능을 비교하였다. FD001 데이터셋에서 test 데이터를 raw RUL 라벨로 평가했을 때, 기존 연구의 Baseline 라벨로 학습한 GRU의 성능은 RMSE 16.459였다(Table 7). 반면 Exponential-Concave 라벨을 적용하면 GRU의 RMSE와 S-Score가 각각 14.515와 301로 나와(Table 8.) 각각 11.81, 22.59%만큼 개선되었다(Table 9). 또한 @CAP=125 평가 조건에서도 RMSE와 S-Score가 각각 12.840과 252로 나타나, 평가 스케일 통일 시에도 일관된 성능 개선 추세를 확인하였다. 전반적으로 제안한 비선형 오목 라벨링 방식은 조각선형 대비 예측 정확도 향상과 지연 벌점(S-Score) 감소를 동시에 달성하였다.

Fig. 5는 FD001 테스트셋에서 임의로 선정한 엔진 유닛의 최종 RUL 예측 결과를 시각화한 것으로, 막대그래프의 주황색(Exp-Concave)은 파란색(True)과 근접하며 초록색(Baseline) 대비 과대/과소 추정이 감소함을 보여준다. 이는 (Table 7-9)

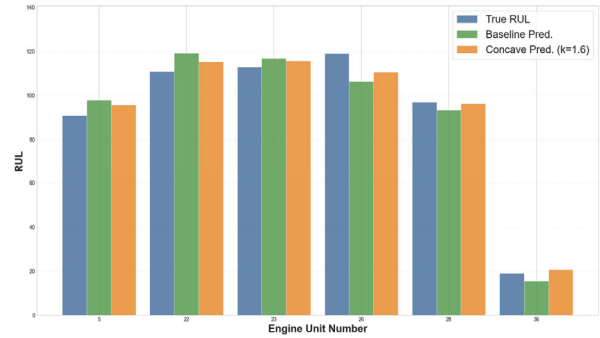


Fig. 5. RUL Prediction Comparison per Engine Unit (Baseline vs Concave)

Table 7. Results with baseline RUL function (CAP=125)

Model	RMSE	MAE	S-Score
GRU	16.459	11.974	388
GRU(@CAP)	15.000	10.904	336
BiLSTM	16.782	11.877	589
BiLSTM(@CAP)	15.560	10.807	546

Table 8. Results with Exp-Concave RUL function (k=1.6)

Model	RMSE	MAE	S-Score
GRU	14.515	10.572	301
GRU(@CAP)	12.840	9.502	252
BiLSTM	16.054	12.055	386
BiLSTM(@CAP)	14.718	10.985	345

Table 9. Performance Improvement of Exp-Concave

Model	Label	RMSE Improvement	S-Score Improvement
GRU	Exp-Concave	11.81%	22.59%
	Exp-Concave(@CAP)	14.40%	25.08%
BiLSTM	Exp-Concave	4.34%	34.53%
	Exp-Concave(@CAP)	5.41%	36.92%

에서 확인된 GRU 기준 RMSE 감소와 S-Score 개선과 일치하며, Exp-Concave 라벨링이 큰 편차를 보이는 RUL 값에 대해 예측 정확도 향상과 늦은 정보 벌점 감소에 기여함을 시각적으로 보여준다.

5. 결론

항공기 터보팬 엔진의 잔여 유효 수명(RUL) 예측은 부품 열화로 인한 고장을 사전에 탐지하여 정비 일정을 최적화하고 안전사고를 예방하는 예지보전(PHM)의 핵심 과제이다. 기존 연구들은 주로 딥러닝 모델 구조의 고도화를 통해 예측 성능

향상을 도모해 왔으나, 학습 목표인 RUL 라벨이 실제 엔진 열화 특성을 충분히 반영하지 못한다는 근본적 한계를 지닌다.

본 연구는 엔진 열화의 비선형적 열역학 특성을 반영하여 최신 연구의 사전학습을 시키는 구조(CAE-RNN-Attention)는 그대로 유지하되, 학습 목표인 RUL 라벨 함수만을 실제 열화 패턴과 유사한 비선형 오목(Concave) 함수로 대체하는 새로운 접근법을 제안한다. 이 방법은 항공기 엔진의 RUL 예측 정확도를 높이기 위해, 기존의 모델 중심적 접근에서 벗어나 학습 목표인 RUL 라벨을 실제 물리적 열화 패턴과 유사한 비선형 함수로 설계하는 새로운 프레임워크를 제안했다.

C-MAPSS FD001 데이터셋을 이용한 실험을 통해, 제안하는 Exponential-Concave 라벨링 방식은 기존의 조각선형 Baseline 라벨링 대비 RUL 예측 오차(RMSE)를 최대 11.81%까지 유의미하게 감소시켰다. 특히, GRU 모델과 $k=1.6$ 의 곡률을 적용한 Exp-Concave 라벨의 조합은 가장 낮은 RMSE인 14.515를 기록하며 제안 방법의 우수성을 보였다. 또한, RMSE를 개선함과 동시에 조기 경보 실패에 더 큰 페널티를 부과하는 안전 지표 NASA S-Score 역시 약 22% 향상시켜, 제안 방식이 정확도와 안전성 모두를 개선하는 효과적인 방법임을 확인하였다.

본 연구는 RUL 예측 문제의 초점을 모델 구조에서 학습 목표 자체의 정교한 설계로 전환하고, 그 효과를 정량적으로 입증했다는 점에서 중요한 기여를 가진다. 향후 연구에서는 제안된 라벨링 프레임워크를 더 복잡한 운용 조건을 가진 다른 데이터셋에 확장 적용하여 적용 가능성과 성능을 더욱 심층적으로 검증할 계획이다.

References

- [1] NASA, C-MAPSS Aircraft Engine Simulator Data [Internet], <https://data.nasa.gov/dataset/c-mapss-aircraft-engine-simulator-data#content>
- [2] D. Frederick, J. DeCastro, and J. Litt, "User's Guide for the Commercial Modular Aero-Propulsion System Simulation (CMAPSS)," *NASA/ARL, Technical Manual TM2007-215026*, 2007.
- [3] S. Russell, P. Norvig, "Artificial Intelligence: A Modern Approach," 3rd ed., New York: Prentice Hall, 2009.
- [4] J. L. Hennessy and D. A. Patterson, "Instruction-level parallelism and its exploitation," in *Computer Architecture: A Quantitative Approach*, 4th ed., San Francisco, CA: Morgan Kaufmann Pub., ch. 2, pp.66-153, 2007.
- [5] D. B. Lenat, "Programming artificial intelligence," in *Understanding Artificial Intelligence*, Scientific American, Ed., New York: Warner Books Inc., pp.23-29, 2002.
- [6] S. Elsherif, B. Hafiz, M. Makhoul, and O. Farouk, "A deep learning-based prognostic approach for predicting turbofan engine degradation and remaining useful life," *Nature Scientific Reports*, 2025.
- [7] A. Heng, S. Zhang, A. C. Tan, and J. Mathew, "Rotating machinery prognostics: State of the art, challenges and opportunities," *Mechanical Systems and Signal Processing*, Vol. 23, pp.724-739, 2009.
- [8] M. Schwabacher, "A survey of data-driven prognostics," in *AIAA Infotech@Aerospace Conference*, 2005.
- [9] A. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol.20, pp.1483-1510, 2006.
- [10] X. Li, Q. Ding, and J. Sun, "Remaining useful life estimation in prognostics using deep convolution neural networks," *Reliability Engineering & System Safety*, Vol.172, pp.1-11, 2018.
- [11] N. Liu, Q. Zhang, and Y. Zhu, "A deep learning approach for remaining useful life estimation using long short-term memory networks," *Mechanical Systems and Signal Processing*, Vol.123, pp.474-486, 2019.
- [12] S. Zheng, K. Ristovski, A. Farahat, and C. Gupta, "Long short-term memory network for remaining useful life estimation," in *Proc. IEEE Int. Conf. Prognostics and Health Management (ICPHM)*, pp.88-95, 2017.
- [13] J. Deutsch and D. He, "Using deep learning-based approach to predict remaining useful life of rotating components," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol.48, No.1, pp.11-20, 2018.
- [14] Y. Zhang et al., "Hybrid CNN-LSTM-attention model for remaining useful life prediction of aero-engines," *Applied Soft Computing*, Vol.109, pp.107503, 2021.
- [15] H. Wang et al., "A hybrid deep neural network model combining CNN, BGRU and self-attention for RUL prediction," *Reliability Engineering & System Safety*, Vol.217, pp.108063, 2022.
- [16] S. Elsherif et al., "A deep learning-based prognostic approach for predicting turbofan engine degradation and remaining useful life," *Scientific Reports*, Vol.15, pp.26251, 2025.



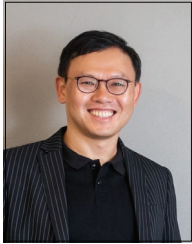
김민정

<https://orcid.org/0009-0002-1286-8037>

e-mail : mjung@sejong.ac.kr

2024년~현재 세종대학교 컴퓨터공학과
학사

관심분야: 인공지능, 예지보전



이강원

<https://orcid.org/0000-0002-3025-4699>

e-mail : kangwon.lee@sejong.ac.kr

1992년 서울대학교 컴퓨터공학과 학사

1994년 서울대학교 컴퓨터공학과 석사

2000년 일리노이 주립대학교 전산학 박사

2000년~2007년 IBM 왓슨 연구소

research staff member

2008년~2014년 IBM 왓슨 연구소 research manager

2014년~2021년 SK 텔레콤 종합기술원 기술원장

2022년~2024년 SK온 DT부문장

2024년~ 현재 세종대학교 컴퓨터공학과 교수

관심분야: 컴퓨터 네트워크, 클라우드 컴퓨팅, 인공지능,
블록체인, 오픈소스 소프트웨어