

Anomaly Audio Signal 활용한
Multi-Output ML/DL 모델 개발



Team Archits

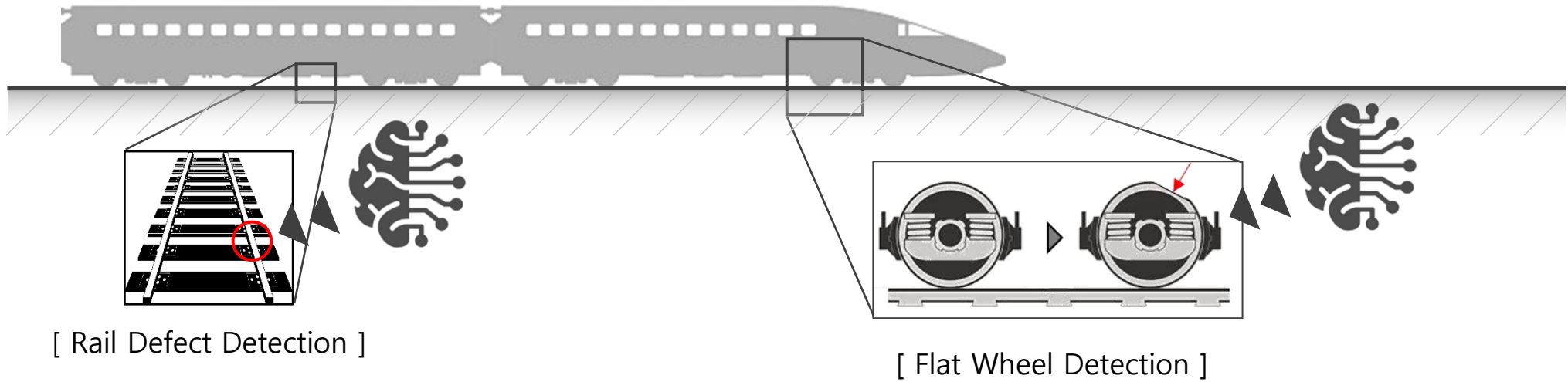
2023.10.20

Contents

- I. 프로젝트 개요
- II. 활용 데이터
- III. 모델 개발 방법
- IV. 실험 및 평가
- V. 활용 계획 및 기대효과
- VI. 시연

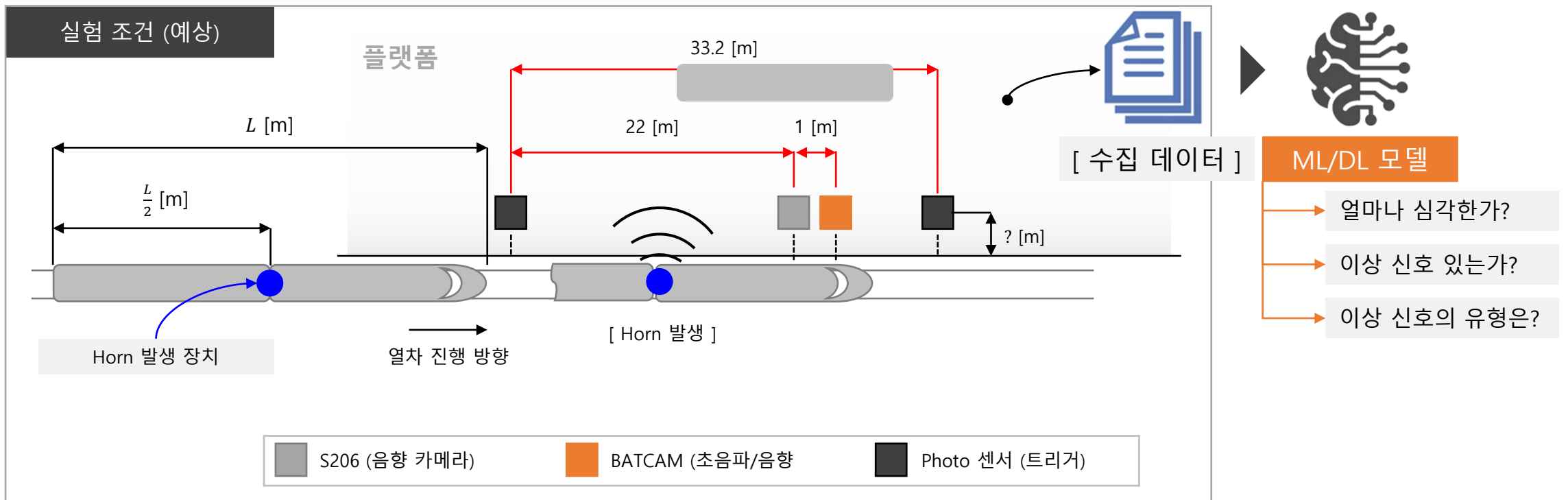
I. 프로젝트 개요

- ▶ 열차는 국가의 중요 인프라이자 향후에도 대체가 어려운 모빌리티임
- ▶ 최근 **AI**를 활용하여 **진동 및 초음파** 기반 대차 Flat wheel, 선로의 Anomaly detection 등에 대한 **연구가 활발**히 진행 중임



I. 프로젝트 개요

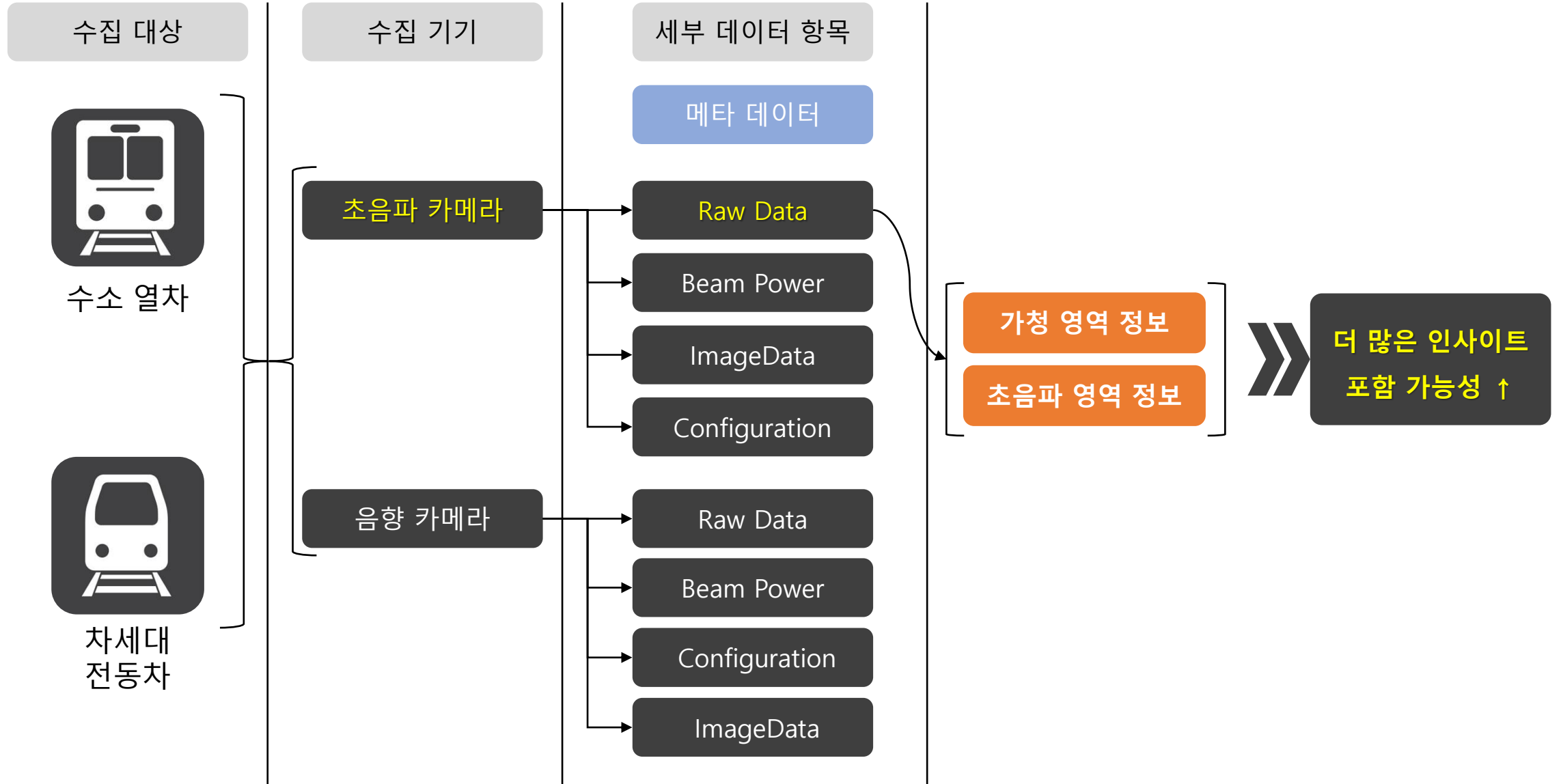
- ▶ Proving Ground에서 실제 운행하는 기차를 대상으로 한 실험 데이터로 판단됨
- ▶ 초음파 카메라와 음향 카메라 등 Audio 데이터 수집하는 과정에서 임의의 시점에 Horn 발생시킴 (Horn 유 = 이상 신호)
- ▶ 예상 최종 성과물을 고려, **Audio 데이터** 활용 **이상 신호** 및 **심각도 판별** ML/DL 모델 구현을 목표로 설정함



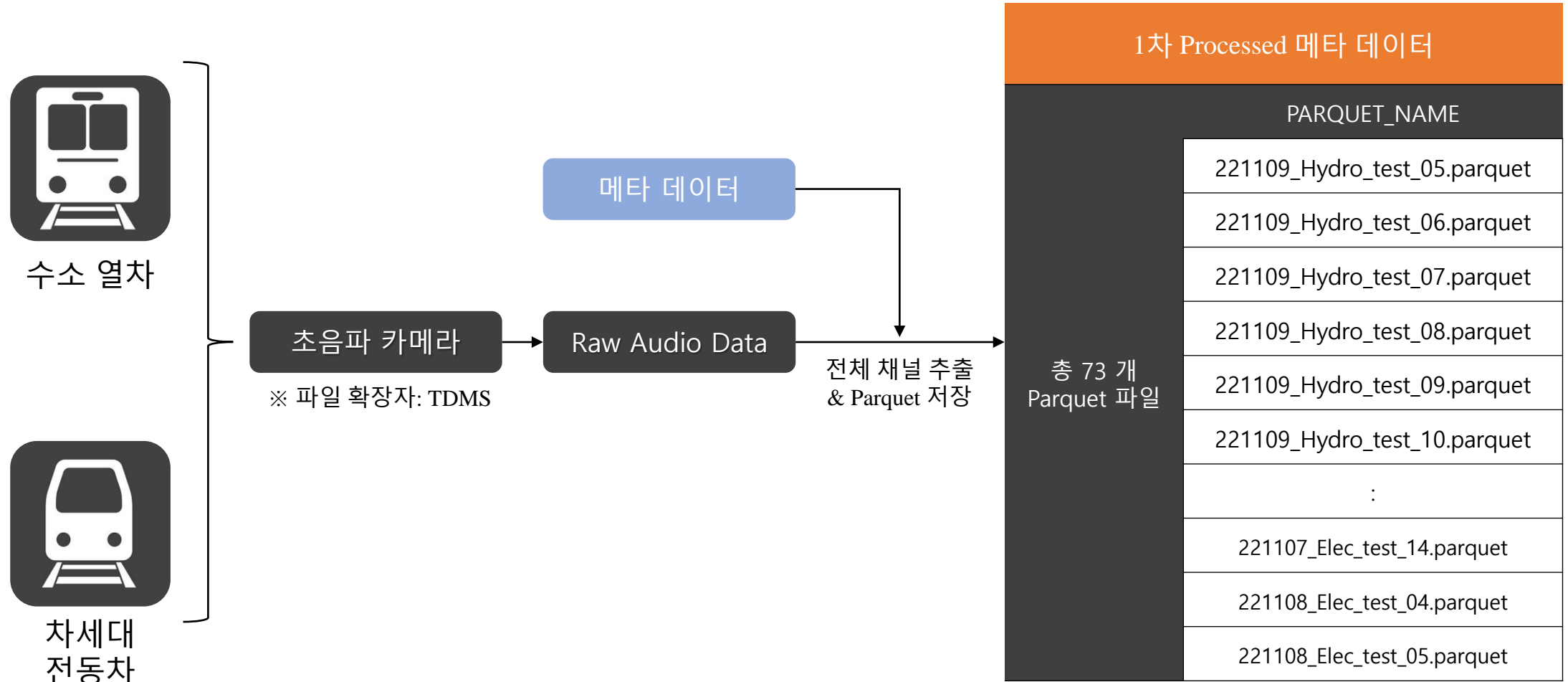


활용 데이터

II. 활용 데이터



II. 활용 데이터 ▶ 데이터 추출 및 저장

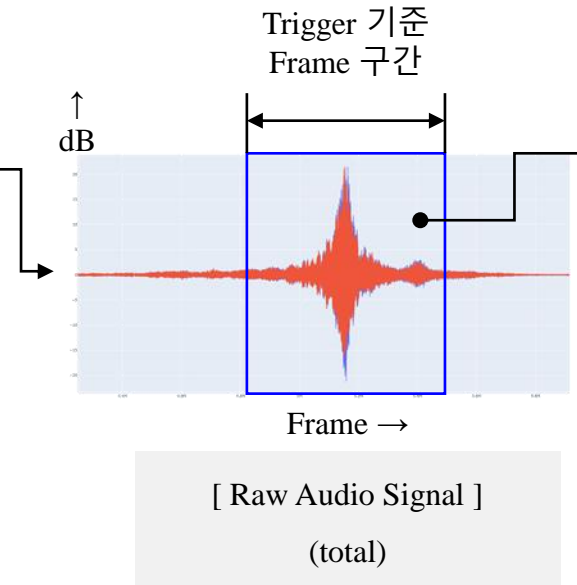


- * TDMS 파일 read 및 데이터 추출: Python 라이브러리 ntms 사용
- * Parquet 파일 생성 저장: Pyarrow 사용

※ 메타데이터와 불일치 혹은 없는 TDMS 제외

II. 활용 데이터 ▶ Horn Data Category 세분화

1차 Processed 메타 데이터
PARQUET_NAME
221109_Hydro_test_05.parquet
221109_Hydro_test_06.parquet
221109_Hydro_test_07.parquet
221109_Hydro_test_08.parquet
221109_Hydro_test_09.parquet
221109_Hydro_test_10.parquet
:
221107_Elec_test_14.parquet
221108_Elec_test_04.parquet
221108_Elec_test_05.parquet



구간 내 Signal Energy 계산

Max Energy 값 & 채널 명 선택

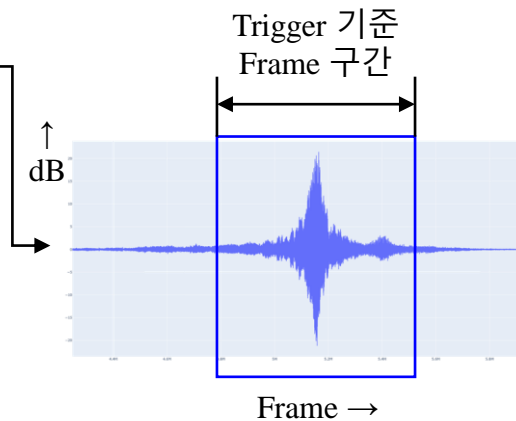
PARQUET_NAME	CHANNEL_NAME	ENERGY
221109_Hydro_test_05.parquet	Channel48	9.74
221109_Hydro_test_06.parquet	Channel48	7.08
221109_Hydro_test_07.parquet	Channel48	7.14
221109_Hydro_test_08.parquet	Channel48	7.41
221109_Hydro_test_09.parquet	Channel94	8.30
221109_Hydro_test_10.parquet	Channel48	8.32
:	:	:
221107_Elec_test_14.parquet	Channel110	563.04
221108_Elec_test_04.parquet	Channel48	413.38
221108_Elec_test_05.parquet	Channel48	850.61

채널 명 추가

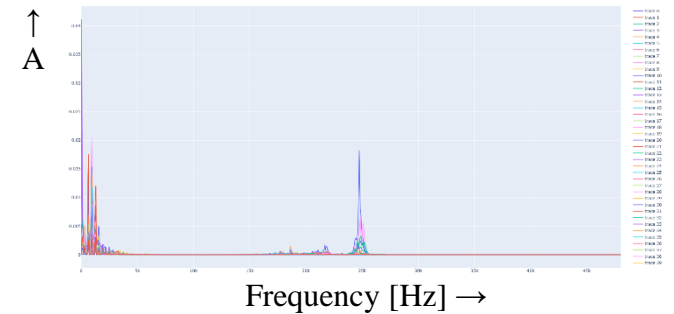
Energy 추가

II. 활용 데이터 ▶ Horn Data Category 세분화

PARQUET_NAME	CHANNEL_NAME	ENERGY
221109_Hydro_test_05.parquet	Channel48	9.74
221109_Hydro_test_06.parquet	Channel48	7.08
221109_Hydro_test_07.parquet	Channel48	7.14
221109_Hydro_test_08.parquet	Channel48	7.41
221109_Hydro_test_09.parquet	Channel94	8.30
221109_Hydro_test_10.parquet	Channel48	8.32
:	:	:
221107_Elec_test_14.parquet	Channel110	563.04
221108_Elec_test_04.parquet	Channel48	413.38
221108_Elec_test_05.parquet	Channel48	850.61



[Raw Audio Signal]
(Channel 48)



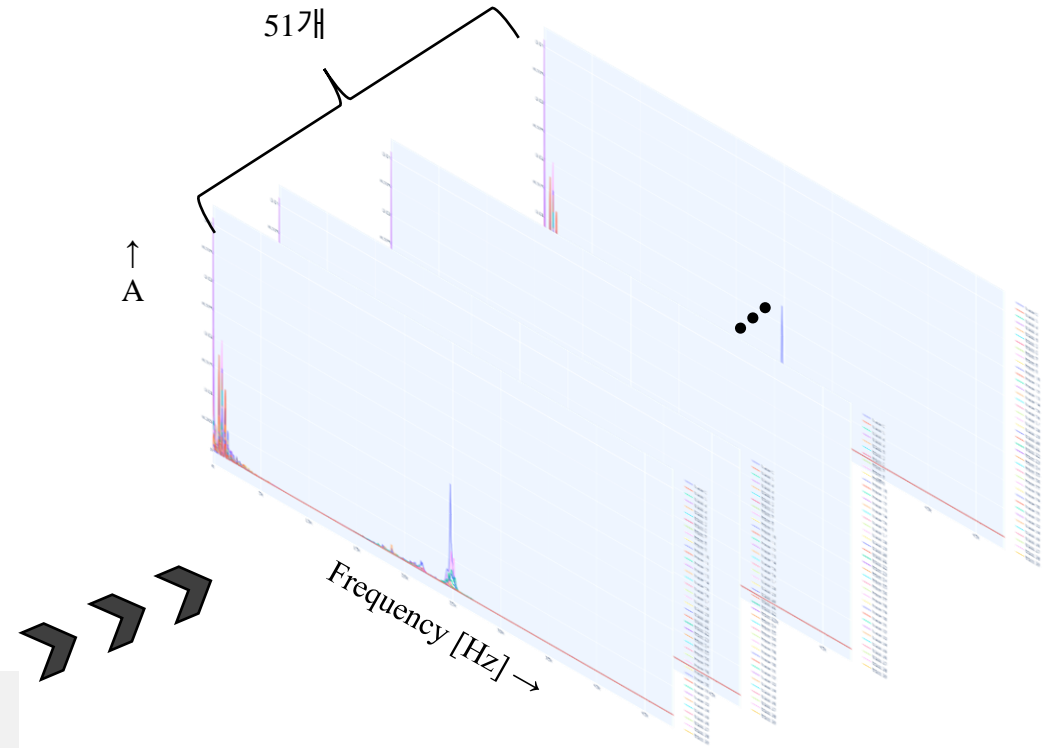
[FFT → Freq. 도메인]
(Max Energy Channel = CHANNEL_NAME)

- * 초음파 카메라 데이터 샘플링 레이트 = 96,000 Hz
- * FFT 라이브러리는 scipy의 welch 함수를 사용하였으며, nperseg = 2**11 (=2048) 로 설정

II. 활용 데이터 ▶ Horn Data Category 세분화

PARQUET_NAME	CHANNEL_NAME	ENERGY
221109_Hydro_test_05.parquet	Channel48	9.74
221109_Hydro_test_06.parquet	Channel48	7.08
221109_Hydro_test_07.parquet	Channel48	7.14
221109_Hydro_test_08.parquet	Channel48	7.41
221109_Hydro_test_09.parquet	Channel94	8.30
221109_Hydro_test_10.parquet	Channel48	8.32
:	:	:
221107_Elec_test_14.parquet	Channel110	563.04
221108_Elec_test_04.parquet	Channel48	413.38
221108_Elec_test_05.parquet	Channel48	850.61

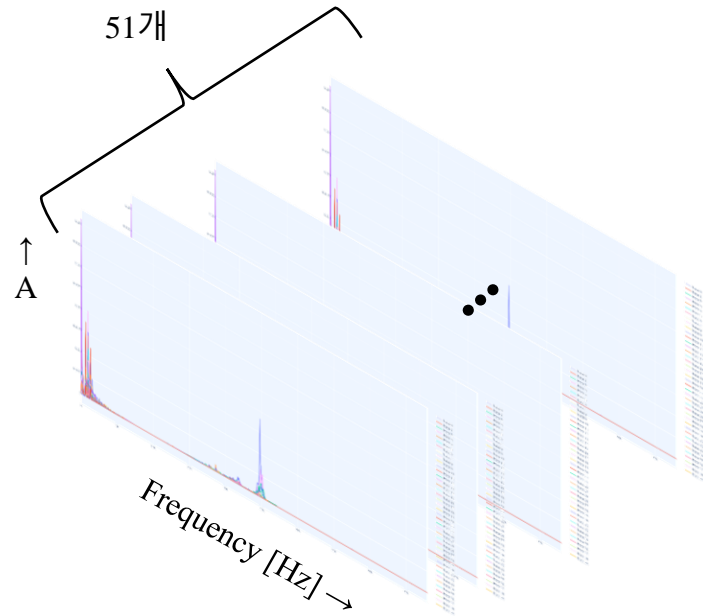
Horn 올린
데이터 51 건



[FFT → Freq. 도메인]
(Max Energy Channel = CHANNEL_NAME)

- * 초음파 카메라 데이터 샘플링 레이트 = 96,000 Hz
- * FFT 라이브러리는 scipy의 welch 함수를 사용하였으며, nperseg = 2**11 (=2048) 로 설정

II. 활용 데이터 ▶ Horn Data Category 세분화

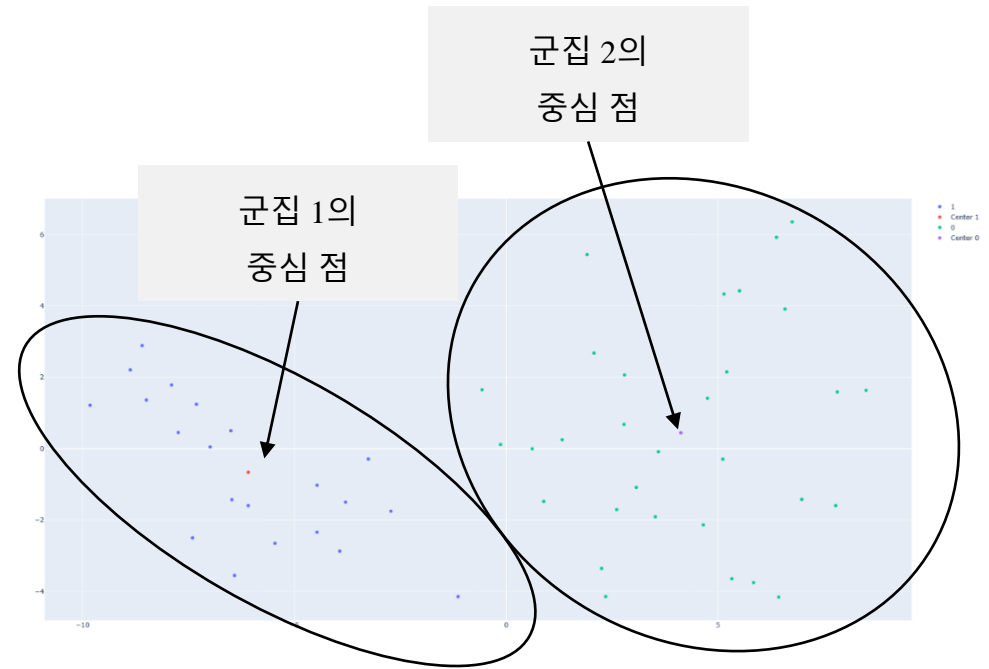


[FFT → Freq. 도메인]
(Max Energy Channel = CHANNEL_NAME)

T-SNE 차원 축소



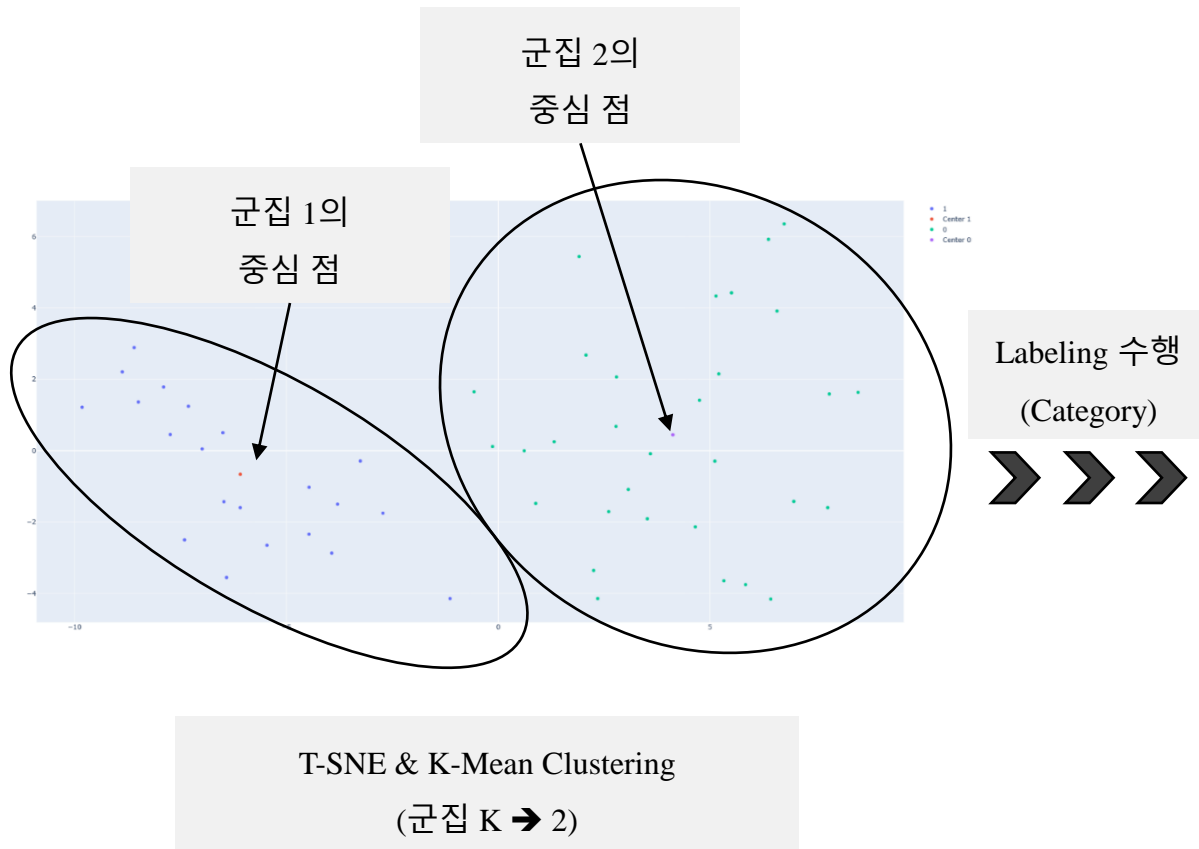
군집화
(K-means Clustering)



T-SNE & K-Mean Clustering
(군집 K → 2)

* T-SNE Parameter = [learning rate : 250, N_iter = 3500, Perplexity = 11]

II. 활용 데이터 ▶ Horn Data Category 세분화



2 nd Processed 메타 데이터			
PARQUET_NAME	CHANNEL_NAME	ENERGY	LABEL
221109_Hydro_test_05.parquet	Channel48	9.74	0
221109_Hydro_test_06.parquet	Channel48	7.08	0
:	:	:	:
221102_Hydro_test_02.parquet	Channel34	0.72	1
221102_Hydro_test_03.parquet	Channel57	10.91	1
221102_Hydro_test_04.parquet	Channel48	16.24	1
:	:	:	:
221107_Elec_test_14.parquet	Channel110	563.04	2
221108_Elec_test_04.parquet	Channel48	413.38	2
221108_Elec_test_05.parquet	Channel48	850.61	2

Horn X

* 해당 카테고리에 대한 물리 현상을 정의할 수 없어서, 1과 2라는 카테고리로 추가 Labeling하여 메타 데이터 업데이트

II. 활용 데이터 ▶ Energy 크기 기반 Severity Labeling

2nd Processed 메타 데이터

PARQUET_NAME	CHANNEL_NAME	ENERGY	LABEL
221109_Hydro_test_05.parquet	Channel48	9.74	0
221109_Hydro_test_06.parquet	Channel48	7.08	0
:	:	:	:
221102_Hydro_test_01.parquet	Channel48	1453.69	2
221103_Elec_test_6.parquet	Channel110	1143.57	1
221103_Elec_test_8.parquet	Channel110	1189.38	1
:	:	:	:
221107_Elec_test_14.parquet	Channel110	563.04	2
221108_Elec_test_04.parquet	Channel48	413.38	2
221108_Elec_test_05.parquet	Channel48	850.61	2



- 1) $E \leq 360$: Severity 0
- 2) $360 < E \leq 600$: 1
- 3) $600 < E \leq 900$: 2
- 4) $E > 900$: 3

Severity (심각도) Labeling

PARQUET_NAME	CHANNEL_NAME	ENERGY	LABEL	SEVERITY
221109_Hydro_test_05.parquet	Channel48	9.74	0	0
221109_Hydro_test_06.parquet	Channel48	7.08	0	0
:	:	:	:	0
221102_Hydro_test_01.parquet	Channel48	1453.69	2	3
221103_Elec_test_6.parquet	Channel110	1143.57	1	3
221103_Elec_test_8.parquet	Channel110	1189.38	1	3
:	:	:	:	:
221107_Elec_test_14.parquet	Channel110	563.04	2	1
221108_Elec_test_04.parquet	Channel48	413.38	2	1
221108_Elec_test_05.parquet	Channel48	850.61	2	2

* Horn 없을 때의 에너지를 Severity 0으로 가정하고, Horn 있을 때 에너지의 각각 25 %tile, 50 %tile, 75 %tile 기준 각각 1, 2, 3 으로 가정

II. 활용 데이터 ▶ 최종 Labeling 메타 데이터

Final Version					
PARQUET_NAME	CHANNEL_NAME	ENERGY	LABEL	SEVERITY	ANOM_POS_METER (이상 소음 위치)
221109_Hydro_test_05.parquet	Channel48	9.74	0	0	0
221109_Hydro_test_06.parquet	Channel48	7.08	0	0	0
221109_Hydro_test_07.parquet	Channel48	7.14	0	0	0
221109_Hydro_test_08.parquet	Channel48	7.41	0	0	0
221109_Hydro_test_09.parquet	Channel94	8.30	0	0	0
221109_Hydro_test_10.parquet	Channel48	8.32	0	0	0
:	:	:	:	:	:
221107_Elec_test_14.parquet	Channel110	563.04	2	1	61
221108_Elec_test_04.parquet	Channel48	413.38	2	1	61
221108_Elec_test_05.parquet	Channel48	850.61	2	2	61

* 이상 소음 위치는 원본 메타 데이터를 활용함

II. 활용 데이터 ▶ 메타 데이터 기반 데이터 셋 구성

- ▶ 모델 학습 목적으로 데이터 셋 구성을 위해 Audio Signal에 대한 Resampling 필요하다 판단함
- ▶ Resampling 길이에 따라 Audio Signal의 Entropy를 계산함으로써 정보의 손실 정도를 파악하고자 함
- ▶ 전반적으로 False Signal Creation 등을 고려 하였을 때, Resampling Frame = 70,000 일 때가 Limit라고 판단함

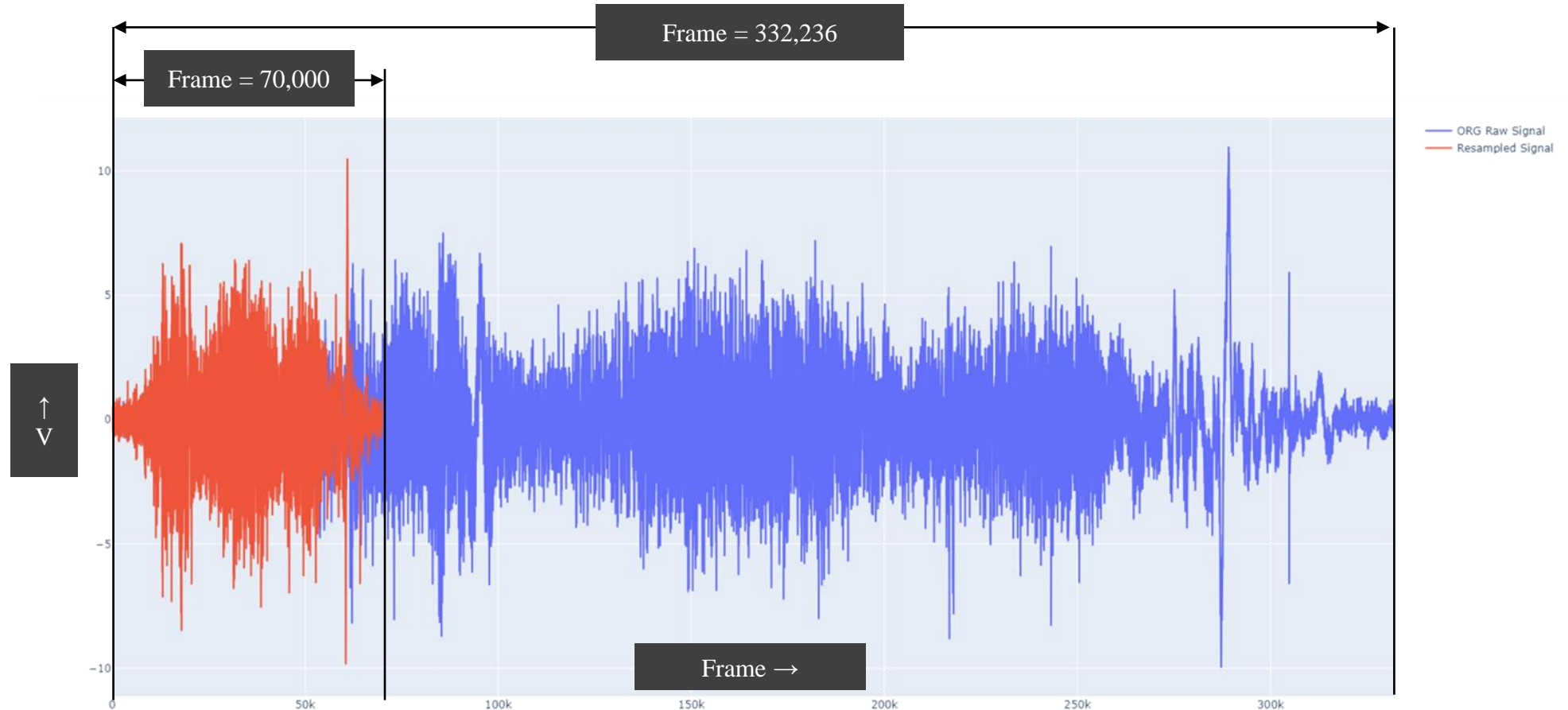
→ False Data Creation

Resampling Value	5000	10000	15000	20000	30000	50000	60000	65000	70000	75000	80000	96000
변환 Channel 수 (전체 파일)	3285	3285	3285	3285	3285	3285	3285	3285	3285	3285	3285	3285
평균 Entropy 변화율	0.39	0.33	0.30	0.28	0.25	0.21	0.19	0.19	0.18	0.18	0.17	0.16
표준편차	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05
최소 Entropy 감소율	0.25	0.19	0.15	0.12	0.09	0.04	0.02	0.01	0.01	0.00 (< 0)	-0.01	-0.02
Entropy 감소율 (25%tile)	0.35	0.29	0.26	0.23	0.20	0.16	0.14	0.14	0.13	0.12	0.12	0.10
Entropy 감소율 (50%tile)	0.40	0.35	0.32	0.30	0.27	0.23	0.22	0.21	0.21	0.20	0.20	0.18
Entropy 감소율 (75%tile)	0.41	0.36	0.33	0.31	0.27	0.24	0.22	0.22	0.21	0.21	0.20	0.19
최대 Entropy 감소율	0.46	0.40	0.38	0.35	0.32	0.29	0.27	0.27	0.26	0.25	0.25	0.24

* 평균 Entropy 감소율: 18.7 ± 4 %

II. 활용 데이터 ▶ 메타 데이터 기반 데이터 셋 구성

- ▶ 모델 학습 목적으로 데이터 셋 구성을 위하여 Raw Audio Signal을 7만 프레임으로 아래와 같이 Resampling 수행함



* 평균 Entropy 감소율: $18.7 \pm 4 \%$

II. 활용 데이터 ▶ 메타 데이터 기반 데이터 셋 구성

	리샘플링 된 Signal														라벨링		위치 예측 위한 데이터			소음 위치		
	0	1	2	3	4	5	6	7	8	9	...	69996	69997	69998	69999	LABEL	SEVERITY	TOT_TIME_SEC	DISTANCE_METER	ANOM_TIME_SEC	ANOMAL_POS_METER	
0	-0.170288	-0.234619	-0.041626	-0.056763	0.030273	0.079468	0.064331	0.109741	0.056763	0.079468	...	0.087036	0.102173	0.113525	0.192993	0	0	1.485510	33.2	0.000000	0.0	
1	-0.151367	-0.155151	-0.083252	-0.018921	0.064331	0.105957	0.087036	0.140015	0.143799	0.079468	...	-0.211914	-0.011353	0.075684	0.007568	0	0	1.485510	33.2	0.000000	0.0	
2	-0.230835	-0.034058	-0.060547	0.075684	0.045410	0.049194	0.071899	0.018921	-0.045410	0.000000	...	0.026489	0.056763	0.018921	0.438965	0	0	1.485510	33.2	0.000000	0.0	
3	-0.128662	-0.143799	-0.011353	0.007568	0.049194	0.068115	0.083252	0.052979	0.000000	0.049194	...	0.158936	0.189209	0.317871	0.238403	0	0	1.485510	33.2	0.000000	0.0	
4	-0.223267	-0.170288	-0.052979	0.052979	-0.022705	-0.007568	0.068115	0.102173	0.034058	-0.075684	...	-0.140015	-0.026489	0.211914	0.234619	0	0	1.485510	33.2	0.000000	0.0	
...
3280	-0.007568	-0.022705	-0.007568	-0.026489	-0.030273	-0.083252	-0.079468	-0.068115	-0.030273	-0.083252	...	0.121094	0.151367	0.155151	0.064331	2	2	3.269026	33.2	6.218031	61.0	
3281	0.045410	-0.011353	0.003784	-0.026489	-0.022705	-0.056763	-0.105957	-0.041626	0.015137	-0.075684	...	0.132446	0.147583	0.181641	0.079468	2	2	3.269026	33.2	6.218031	61.0	
3282	0.034058	-0.003784	0.022705	-0.003784	-0.049194	-0.056763	-0.075684	-0.052979	-0.018921	-0.075684	...	0.124878	0.132446	0.158936	0.124878	2	2	3.269026	33.2	6.218031	61.0	
3283	0.003784	-0.015137	-0.007568	0.000000	-0.034058	-0.071899	-0.052979	-0.064331	-0.003784	-0.083252	...	0.121094	0.166504	0.166504	0.105957	2	2	3.269026	33.2	6.218031	61.0	
3284	0.060547	-0.007568	0.022705	0.003784	-0.041626	-0.068115	-0.071899	-0.056763	-0.022705	-0.056763	...	0.132446	0.181641	0.162720	0.075684	2	2	3.269026	33.2	6.218031	61.0	

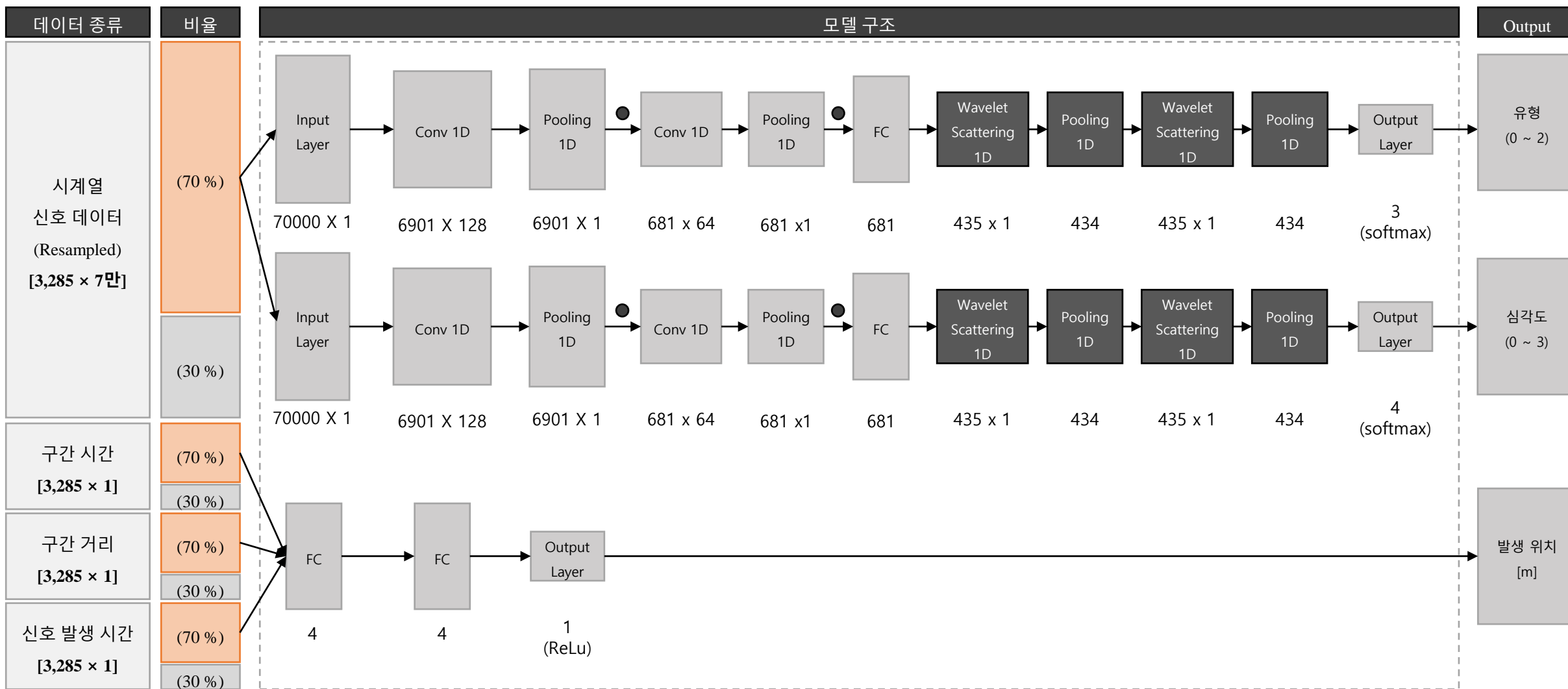
3285 rows × 70006 columns

III. 모델 개발 방법

모델 1) Wavelet Scattering Transform + MLP 기반 MIMO 모델

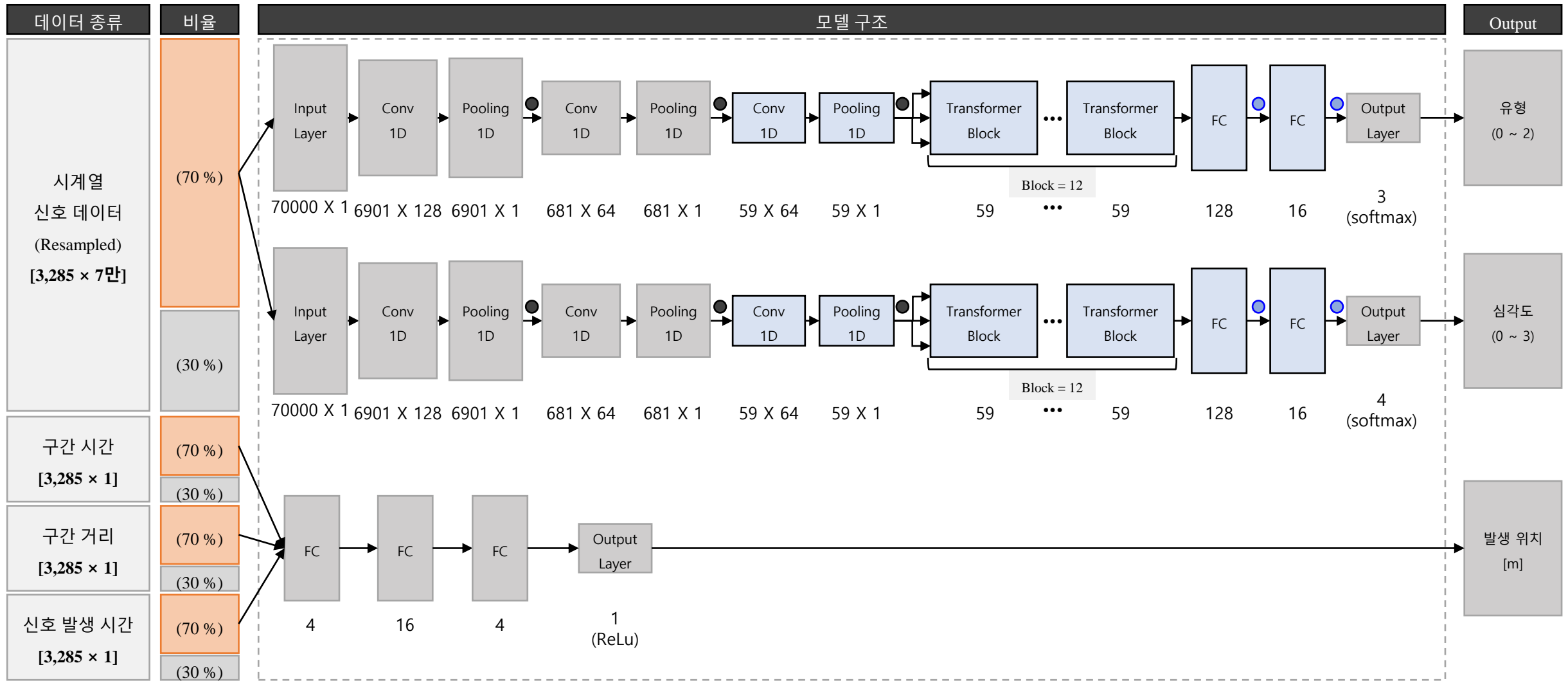
모델 2) Transformer + MLP 기반 MIMO 모델

III. 모델 개발 방법 ▶ MIMO : WST + MLP



* ● (Drop out) = 0.4

III. 모델 개발 방법 ▶ MIMO : Transformer + MLP

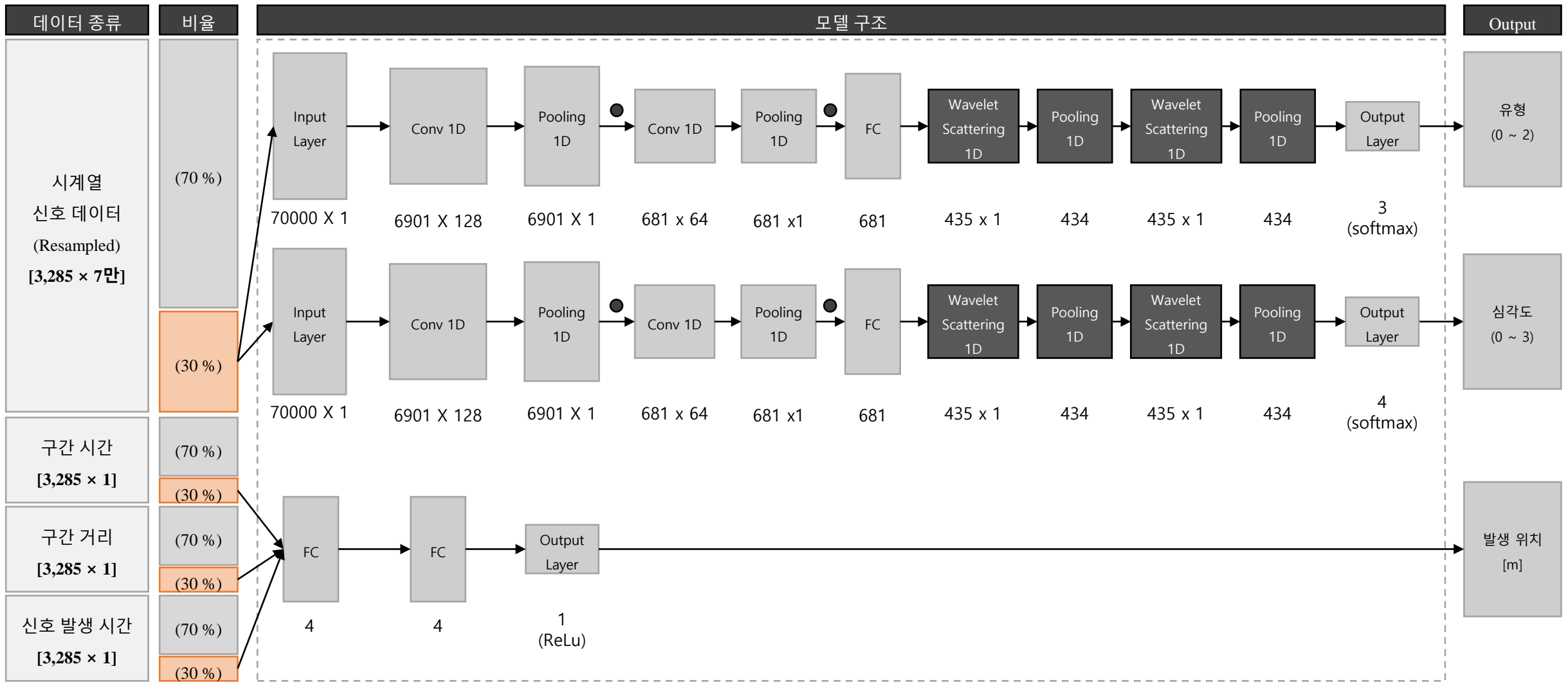


* ● (Drop out) = 0.3, ● (Drop out) = 0.3



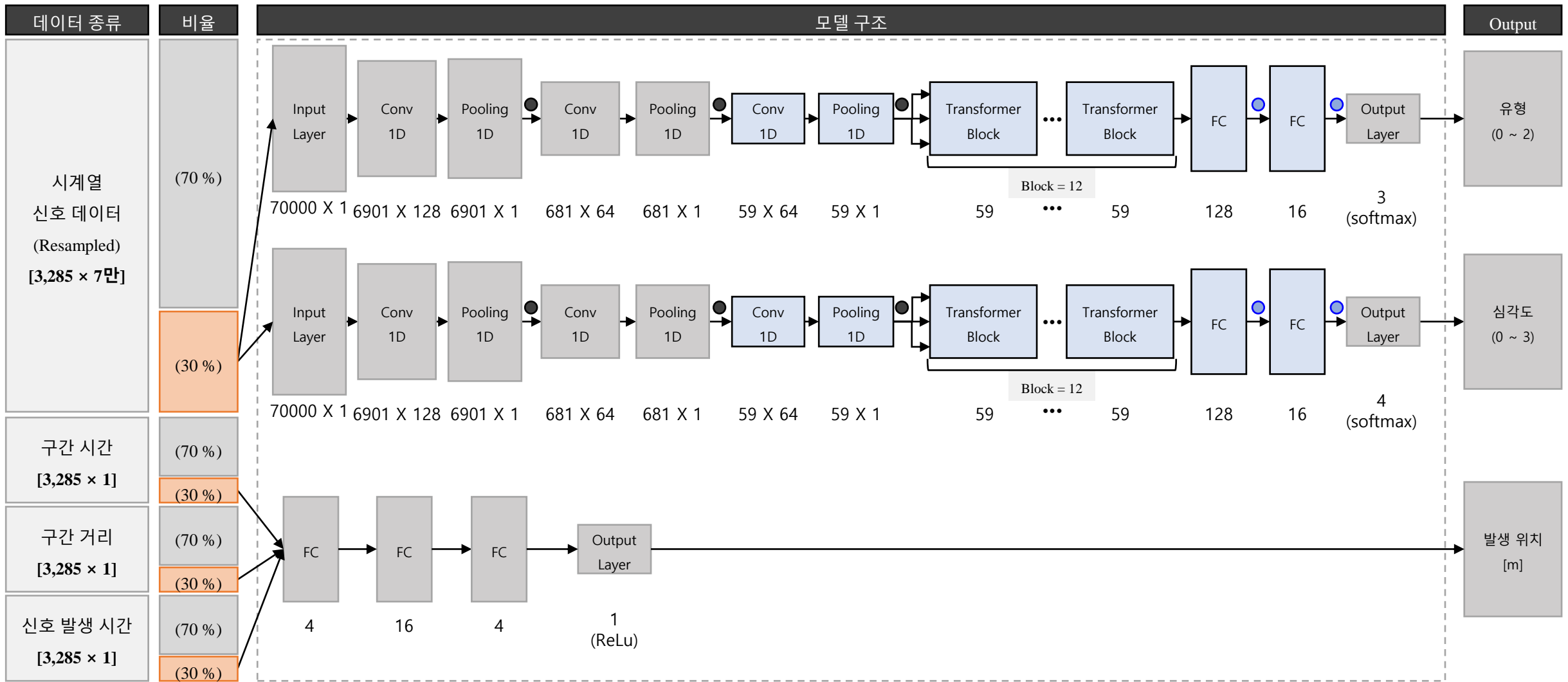
실험 및 모델 평가

IV. 실험 및 평가 ▶ MIMO : WST + MLP



* ● (Drop out) = 0.4

IV. 실험 및 평가 ▶ MIMO : Transformer + MLP



* ● (Drop out) = 0.3, ● (Drop out) = 0.3

IV. 실험 및 평가 ▶ 비교 평가 결과

- ▶ 각 주제별 **분류 정밀도**, **신뢰도**, 테스트 데이터 기준 **평균 소요시간**을 고려하였을 때 **Transformer 모델로 선정함**
- ▶ 그러나 소음 위치 예측의 경우, 두 모델 모두 데이터가 너무 단순하여 모델이 학습할 패턴을 찾지 못한 것으로 판단됨

분류/예측 대상	모델 타입	Classification			Prediction			Avg. Elapsed Time [s]
		Precision [%]	Recall [%]	F1 Score	MAPE [%]	MAE [m]	RMSE [m]	
Category (이상 신호 유형)	WST	99.9	99.9	0.999				0.435
	Transformer	99.8	99.8	0.998	-	-	-	0.197 (~ 221% ↑)
Severity (심각도)	WST	79.4	84.7	0.807				0.435
	Transformer	95.6 (20% ↑)	95.1 (12% ↑)	0.951	-	-	-	0.197
Position (소음 위치)	MLP	-	-	-	X	32.2	1997.3	0.435
	MLP	-	-	-	X	32.2	1997.3	0.197

V. 활용 계획 및 기대효과

기대 효과

- ✓ Audio Signal 기반 **Transformer Model 성능**이 WST 기반 모델에 비하여 **우수하다** 판단됨
- ✓ 현재 모델의 구조 특성상 **초음파 Audio 센서 한 개**의 데이터로도 운용 가능함
- ✓ 실시간 Audio 수집 **Edge device** 및 **AR 기기** 추가 → **현장 승무원**에게 이상 상태의 종류, 심각도 등을 **실시간 정보 제공** 예상

추가 필요 사항

- ✓ **Hyper Parameter Tuning** (Grid searching) 필요
- ✓ **위치 예측** 패턴 학습 관련 **학습 파라미터 추가 필요** (e.g., DoA 등)

활용 계획

- ✓ 추가 성능 검증 가능 여부 확인 및 본 프로젝트 결과를 기반으로 논문 작성
- ✓ 성능 검증 후 AR 환경에서 파악할 수 있도록 시스템 구성

The image features a minimalist design with a white background. At the top, there are three horizontal bars of varying lengths and shades of gray, extending from the right edge towards the left. At the bottom, there are three similar horizontal bars of varying lengths and shades of gray, extending from the left edge towards the right. The word "Fin" is centered in the middle of the page in a black, serif font.

Fin