

최연희^{1,3}, 하은택^{2,3}, 안병현^{1,3}, 김민재⁴, *백준기^{4,5}

{gregoyo, het117, anpos1}@naver.com, {imkbsz, paikj}@cau.ac.kr

대한민국 육군¹, 해군²

중앙대학교 국방 AI대학 핵심인재과정³, 영상학과⁴, AI학과⁵

연구 목적 및 데이터

- 전차, 장갑차, 각종 전투차량 등 총 23종의 **군 운송수단**에서 발생하는 **소리를 구분하는** 인공지능 모델 설계
- AI-Hub^[1]의 「자연 및 인공적 발생 **非 언어적 소리 데이터**」에서 제공하는 2,369개의 군 운송수단 음향 신호를 활용
- 데이터는 6:2:2로 나누어 각각 학습:검증:평가로 이용

구분	군 운송수단 음향 신호원 (23종)
전차 (2)	K-1, K-1a1
궤도·장갑차 (10)	K-56, K-77, K288a1, K-200, K800, 화생방정찰차, K10 탄약 운반차, Km9ace, 교량 전차, 장애물개척전차
차륜 전투차량 (11)	2.5t, 9.5t, 5t, 10t, 27t, 다목적 굴착기, 살수차, 대형버스, 부식수송차량, 승용차, 통신 가설 차량

실험 결과

- 실험 환경

구현	TensorFlow, Keras
GPU	Nvidia Geforce RTX 3070 Laptop

- 모델 및 전처리 방법에 따른 분류 정확도

Model		Accuracy (%)	Model		Accuracy (%)
LR	ZCR	20.04	LeNet5	MFCC	83.33
	Centroid	14.56		Chroma	53.8
	Roll-Off	17.09		Mel	43.67
	MFCC	86.92		ZCR	21.1
	Chroma	44.3		Centroid	14.56
	Mel	56.96		Roll-Off	24.05
SVM	ZCR	14.56	VGG16	MFCC	67.72
	Centroid	16.03		Chroma	42.62
	Roll-Off	17.3		Mel	54.01
	MFCC	87.55		ZCR	25.74
	Chroma	43.46		Centroid	21.73
	Mel	61.81		Roll-Off	25.53
LeNet5	ZCR	22.78	ResNet50	MFCC	76.79
	Centroid	21.52		Chroma	51.69
	Roll-Off	23.84		Mel	48.87

- Epoch 설정에 따른 분류 정확도

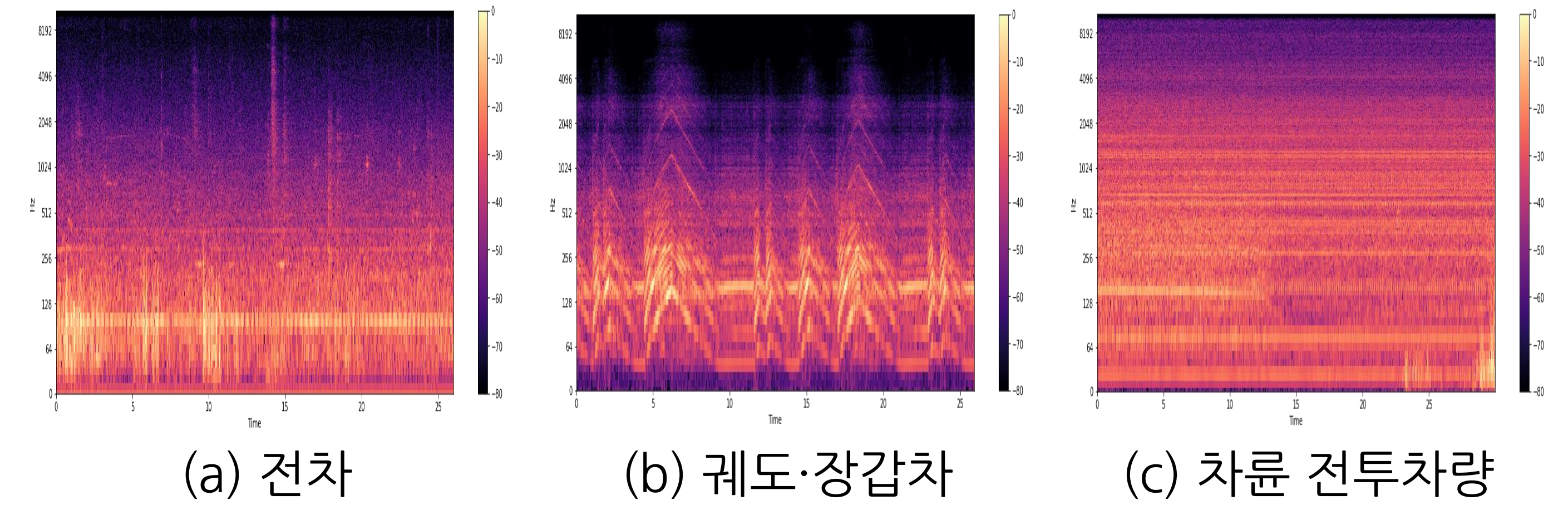
Model	LR		SVM		LeNet5		VGG16		ResNet50	
Epoch	100	300	100	300	100	300	100	300	100	300
Accuracy (%)	84.2	86.9	83.5	87.6	83.3	82.1	62.2	67.7	76.8	71.3

- 모델별 평균 학습 및 평가 시간 (sample/second)

Model	LR		SVM		LeNet5		VGG16		ResNet50	
Train	0.0237	0.0229	0.0637	3.8202	4.3865					
Test	0.0004	0.0004	0.0007	0.0124	0.0065					

제안하는 방법

- 시간-주파수 표현을 이용한 군 운송수단 주파수 특성 분석
↳ **고주파 영역**을 중심으로 **유의미한 차이** 발견

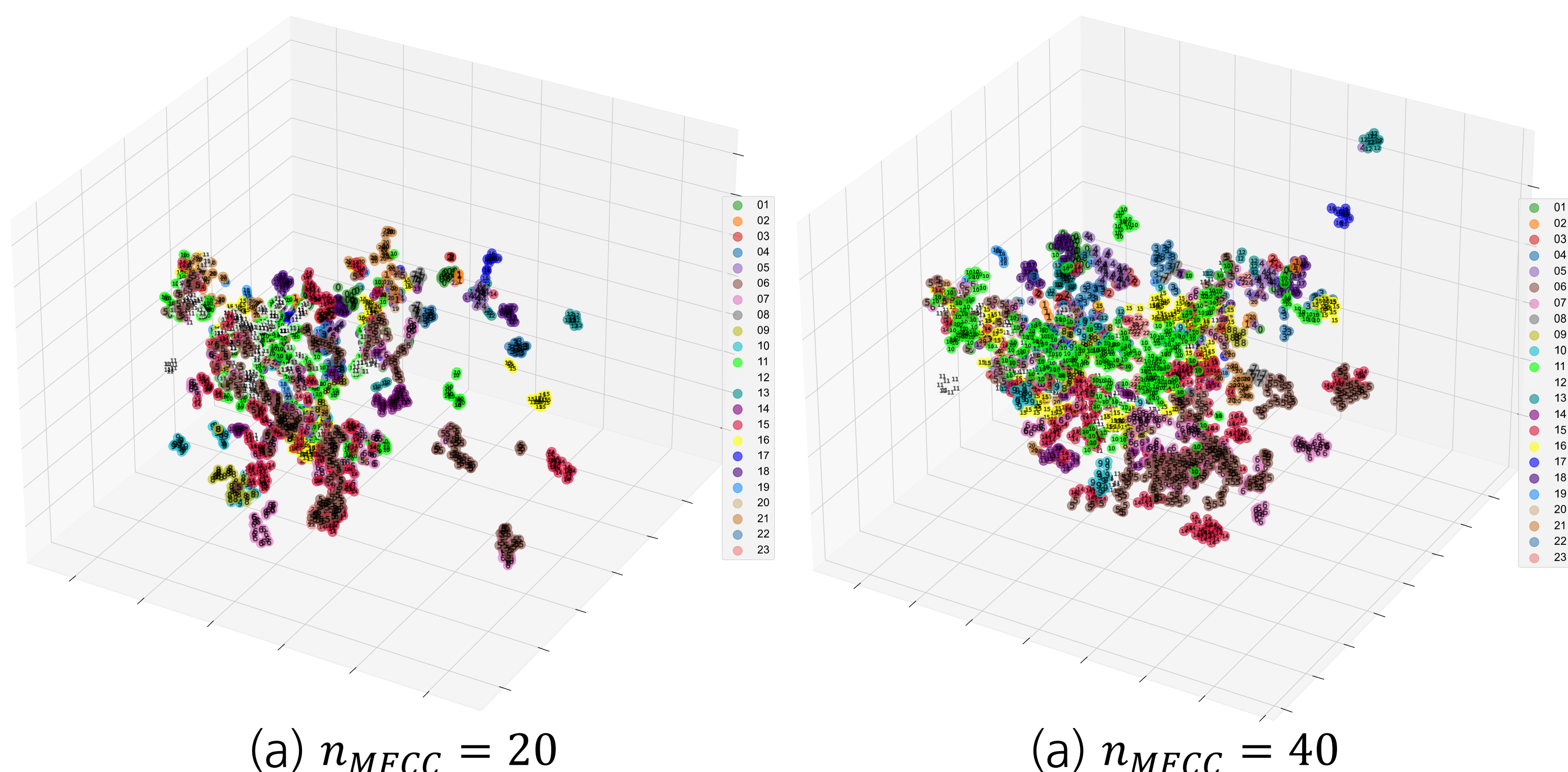


- 주파수 특성을 활용하기 위한 **특징 추출** 방법 선정:
(1) Zero-crossing rate, (2) Spectral centroid, (3) Spectral roll-off, (4) MFCC, (5) Chroma frequencies, (6) Mel spectrogram
- 다양한 복잡도를 갖는 **머신/딥 러닝** 방법 선정:
(1) Linear regression, (2) SVM, (3) LeNet, (4) VGG, (5) ResNet
- 특징 추출 방법 (6) ✗ 인공지능 모델 (5) ✗ 학습조건 (4)
↳ 120개 조합으로 비교 실험 수행

특징 추출 방법	모델 종류	학습 조건	
Zero-crossing rate	LR	Optimizer	Adam
Spectral centroid	SVM		SGD
Spectral roll-off ✗	LeNet5 ✗	Epoch	100
MFCC	VGG16		300
Chroma frequencies	ResNet50		
Mel spectrogram			

결론 및 고찰

- 적절한 특징 추출 과정이 수반되는 경우 (**MFCC**), 휴대용 단말기에서 구동이 가능할 정도로 경량인 선형 모델로 80%이상의 정확도로 군 운송 수단을 분류할 수 있음을 확인
- 음향 신호 MFCC 특징 벡터 t-SNE^[2] 시각화
↳ 고주파 성분이 충분히 보존된 상황에서 군 운송수단의 음향 신호는 선형 분리가 가능



- 향후 연구에서는 딥러닝 모델의 성능을 이끌어내기 위해 데이터 증강, 정칙화, 사전학습 등의 방법 등을 시도할 계획

[1] AIHub Homepage

Available online: <http://www.aihub.or.kr/> (accessed on 31 August 2022).

[2] t-SNE

L. van der Maaten and G.E. Hinton, "Visualizing High-Dimensional Data Using t-SNE," J. Machine Learning Research, vol. 9, pp. 2579-2605, 2008.