

전문가 지식을 활용한 군집화를 통한 확산 모델 기반 차량 내 소음 증강

(Diffusion Model-based In-vehicle Noise Augmentation through Expert Knowledge-guided Clustering)

최 석 훈[†] 백 무 근[†] 부 석 준^{**}
(Seok-Hun Choi) (Mugeun Baek) (Seok-Jun Buu)

요약 차량 운행의 안전성과 사용자 경험 향상을 위해, 차량 내 소음을 정확히 분류하고 이를 기반으로 성능 이상을 사전에 예측하는 기술이 중요하다. 그러나 기존 딥러닝 기반 분류기는 외부 잡음, 내부 반향음 등 복잡한 음향 환경에서 성능 저하를 보인다. 본 논문에서는 이러한 문제를 해결하기 위해, 확산 모델 기반 신호 증강으로 복잡한 차량 내 음향 환경을 모사하고, 신호-라벨 간 상관관계를 강화하는 데이터 증강 기법과 전문가 지식 기반 군집화 기법을 결합한 새로운 차량 소음 분류 기법을 제안한다. 제안한 방법은 다양한 소음 조건을 반영한 합성 데이터를 생성하고, 전문가 정의에 따른 군집화 기반 라벨링을 통해 실제 환경을 반영한 학습 데이터를 구성함으로써 분류 성능을 향상시킨다. 실험 결과, 본 방법은 최신 딥러닝 기반 분류기 대비 99.60%의 정확도를 달성했으며, 기존 생성 모델 기반 증강보다 0.06%p 높은 성능을 보여 그 효용성을 입증하였다.

키워드: 확산 모델, 생성적 딥러닝, 데이터 증강, 음향 분류, 차량 소음 분류

Abstract To ensure vehicle operational safety and enhance user experience, it is crucial to accurately classify in-vehicle noise and detect performance anomalies in advance. However, deep learning-based noise classifiers often struggle in complex acoustic environments, such as those with external noise and internal reverberation. To address these challenges, we propose a novel vehicle noise classification method that integrates diffusion model-based signal augmentation with expert knowledge-guided clustering. This approach synthesizes a variety of challenging in-vehicle acoustic conditions and enhances signal-label associations through automatic label assignment based on expert-defined clusters. As a result, we can create training datasets that closely mirror real-world scenarios. Our experiments demonstrate that this method achieves a classification accuracy of 99.60%, surpassing state-of-the-art classifiers and improving by 0.06 percentage points over existing generative augmentation methods, thereby showcasing the effectiveness of the diffusion-based approach.

Keywords: diffusion models, generative deep learning, data augmentation, acoustic classification, in-vehicle noise classification

· 이 논문은 (주)현대자동차와 (주)에스엠인스트루먼트의 지원을 받아 수행된 연구의 결과물임

· 이 논문은 2024 한국컴퓨터종합학술대회에서 '차량 내 소음 분류를 위한 의사 라벨링 알고리즘을 사용한 확산 모델 기반 신호 증강'의 제목으로 발표된 논문을 확장한 것임

[†] 학생회원 : 경상국립대학교 컴퓨터공학부 학생
hse110025@gnu.ac.kr
arcade123@gnu.ac.kr

^{**} 정 회 원 : 경상국립대학교 컴퓨터공학부 교수/항공우주공학부 JA 교수
(Gyeongsang Nat'l Univ.)
sj.buu@gnu.ac.kr
(Corresponding author)

논문접수 : 2024년 9월 25일
(Received 25 September 2024)

논문수정 : 2025년 5월 13일
(Revised 13 May 2025)

심사완료 : 2025년 7월 11일
(Accepted 11 July 2025)

Copyright©2025 한국정보과학회 : 개인 목적이거나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.
정보과학회논문지 제52권 제9호(2025. 9)

1. 서론

차량 내 소음의 정확한 분류는 차량의 성능을 진단하고 결함을 파악해서 안전성을 향상하는 중요한 일이다. 그리고 이는 사용자 경험 개선에 핵심적인 역할을 한다[1]. 차량에서 발생하는 다양한 소음은 단순한 기계적 작동의 결과일 뿐만 아니라, 도로 상태, 운전 습관, 주변 환경 조건 등 외부 요인들에 의해 크게 영향을 받는다. 이러한 복잡한 요인들이 결합하여 차량 내부에서 다양한 음향 특성이 발생하며, 이는 시간에 따라 변동하거나 불규칙하게 나타나기 때문에 이를 효과적으로 분류하는 것은 매우 어려운 과제이다[2,3]. 기존의 분류 방법들은 이러한 배경 소음의 다양성과 차량 내부의 반향, 굴절 등으로 인해 소음의 복잡한 특성을 모두 반영하지 못해 제한된 성능을 보인다[3]. 복잡한 차량 내부 환경에서의 소음 분류를 위해서는 다변량 소음의 패턴 추출과 데이터의 다양성을 극대화하는 한 단계 높은 단계의 신호 증강 방법이 필요하다.

본 논문에서는 기존 소음 분류 기법들이 복잡한 차량 내 음향 환경을 충분히 반영하지 못해 성능에 한계를 보이는 문제를 해결하기 위해, 효과적인 데이터 증강과 분류 성능 향상을 목표로 한 새로운 접근법을 제안한다. 제안하는 방법의 주요 공헌은 확산 모델 기반 신호 증강과 전문가 지식 기반 군집화 기법을 결합하여 복잡한 소음 패턴을 보다 정확하게 반영할 수 있다는 점이다. 이 방법에서는 차량 내 소음 데이터가 확산 모델 기반 신호 증강 기법을 통해 생성된다. 그리고 생성된 데이터에 대해 전문가 지식 기반 군집화 과정을 거쳐 복잡한 데이터를 군집화 한다. 이렇게 증강된 데이터를 확보하여 분류기의 성능을 극대화한다. 이로써, 기존의 단순한 증강 방법들보다 더 복잡하고 다양한 소음 조건을 반영하면서도 고품질의 합성 소음 샘플을 생성할 수 있다. 본 논문에서 제시한 확산 모델 기반 신호 증강 방법은 전통적인 증강 방법과 딥러닝을 활용한 다른 증강 방법들에 비해 최소 0.06%p의 성능 향상이 있었으며, 이는 제안하는 방법이 더 복잡하고 다양한 소음 패턴을 학습하는 데 효과적임을 시사한다.

2. 관련연구

표 1에서는 음향 데이터셋의 분류 정확도를 개선하기 위해 전통적인 증강 방법과 딥러닝 기반의 새로운 증강 방법을 적용하는 시도를 정리하였다. Flip, Rotate, Blur, Shift 등의 전통적인 증강 방법은 주로 간단한 기하학적 변형을 통해 데이터의 다양성을 높이는 데 초점을 맞추고 있다. 이러한 방법들은 특히 데이터셋의 크기가 상대

적으로 작거나 특정 패턴에 과적합되는 것을 방지하는데 유용하다[4]. 예를 들어, Stanford Cars-196 데이터셋에서 이러한 전통적인 증강 방법들을 적용했을 때, 분류 정확도가 94.60%로 향상되었으며, 이는 이들 방법이 데이터의 여러 특성을 모델이 학습할 수 있도록 도와준 결과라 할 수 있다[4,5].

딥러닝 기반의 새로운 증강 방법들은 전통적인 방법들에 비해 더 복잡한 데이터 변형을 수행한다. GAN, ACGAN, DCGAN 및 DEVAE-GAN과 같은 생성적 적대 신경망(GAN) 기반 방법들은 실제 데이터와 유사한 새로운 데이터를 생성하여 모델의 학습에 사용된다. 이러한 방법들은 기하학적 변형을 통한 전통적인 증강기법보다 복잡한 신호 패턴을 포함하는 데이터셋에서 유용하며, 다양한 합성 데이터의 생성을 통해 모델이 다양한 상황을 학습할 수 있도록 한다. 이를 통해 UrbanSound8K, In-vehicle noise, EEG-based SEED 등의 다양한 데이터셋에서 각각 97.03%, 95.81%, 97.21%의 높은 분류 정확도를 기록하였다[6-9].

기존의 데이터 증강 기법들은 차량 내 환경처럼 다양한 소음원이 혼재된 복잡하고 변칙적인 음향 조건에서는 효과가 제한적이다. 단순한 기하학적 변형은 시간-주파수 도메인의 복잡한 패턴을 충분히 반영하지 못하며, GAN 기반 생성 기법은 훈련의 불안정성과 해석 가능성 부족으로 인해 실제 적용에 어려움이 따른다[10]. 특히, 생성된 데이터가 차량 소음처럼 고차원적이고 비선형적인 특성을 충분히 반영하지 못하는 경우가 많아, 기존 증강 방법은 이러한 변칙적 상황에서 성능 저하를 겪는다. 이러한 한계를 극복하기 위해 본 논문은 확산 모델 기반의 고성능 신호 증강 기법을 제안한다. 확산 모델은 초기 잡음 상태에서 점진적으로 신호를 복원함으로써, 차량 소음의 미세하고 복합적인 패턴을 정밀하게 학습할 수 있다[10].

최근 연구들에서도 확산 모델이 음성 합성, 음향 향상, 이상 음향 탐지 등 다양한 분야에서 기존 생성 모델을 능가하는 성능을 보이며, 안정적으로 복잡한 시간-주파수 구조를 학습할 수 있음을 입증하고 있다[11-13]. 본 연구는 이러한 확산 모델의 특성을 차량 내 소음 분류를 위한 데이터 증강에 적용하고, 여기에 전문가 지식 기반의 군집화 기법을 결합하여 증강 데이터에 대한 라벨링 정확성과 효율성을 동시에 확보하는 새로운 접근을 제안한다. 이를 통해 시간과 자원의 소모를 줄이면서도 높은 품질의 증강 데이터를 확보할 수 있으며, 복잡한 실환경에서도 효과적으로 작동하는 데이터 증강 전략을 제시한다.

표 1 음향 데이터 및 증강 방법에 따른 분류 연구
Table 1 Related Work Utilizing URL and HTML for Phishing Detection

Augmentation Type	Year	Method	Dataset	Classification Performance
Traditional Augmentation	2018	Flip, Rotate, Blur, Noise[4]	Stanford Cars-196	ACC: 0.9460
	2023	Flip, Shift[5]	Exhaust system vehicle	ACC: 0.9840
Deep Learning-based Augmentation	2019	GAN[6]	UrbanSound8K	ACC: 0.9703
	2021	ACGAN[7]	In-vehicle noise	ACC: 0.9581
	2022	DCGAN[8]	UrbanSound8K	ACC: 0.9800
	2023	DEVAE-GAN[9]	EEG-based SEED	ACC: 0.9721

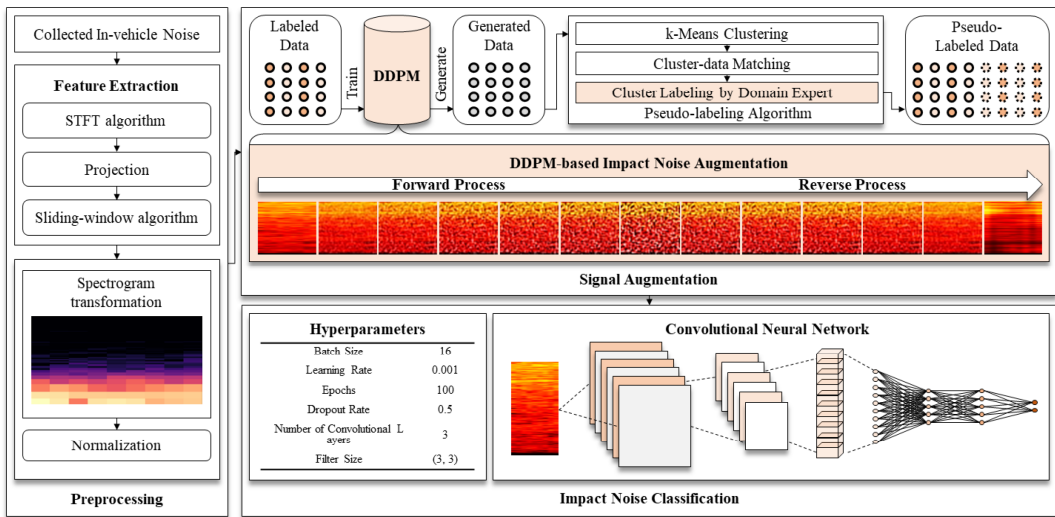


그림 1 차량 내 소음 분류를 위한 전체 구조

Fig. 1 Autoencoder-based Anomaly Detection and Multimodal Ensemble Method

3. 제안하는 방법

본 연구에서는 총 10개 클래스로 구성된 차량 내 소음 데이터를 30채널 센서 배열을 통해 수집하고, 44.1kHz로 샘플링하였다. 수집된 오디오 신호는 STFT(Hanning 윈도우, 1024-point, hop size 256)를 이용해 시간-주파수 도메인의 스펙트로그램으로 변환되며, 1초 길이의 고정된 윈도우로 분할하여 정규화 후 모델에 입력된다.

확산 모델은 DDPM 구조를 기반으로 하며, 1000 step의 노이즈 스케줄을 따라 학습된다. 학습된 모델은 denoising 및 pseudo-labeling 과정을 통해 새로운 합성 샘플을 생성한다. 이후 SSIM 기반 필터링과 KMeans 기반 근집화를 통해 전문가 지식에 따라 라벨을 할당하고, 기존 데이터와 함께 CNN 분류기의 학습에 사용된다.

이러한 처리 과정은 복잡한 차량 내 소음을 효과적으로 반영하며, 라벨링 비용을 줄이고 증강 데이터의 품질

을 향상시키는 데 기여한다.

3.1. 단시간 푸리에 변환과 슬라이딩 윈도우 알고리즘

단시간 푸리에 변환(STFT)은 시간-주파수 분석을 위해 신호에 대해 슬라이드 하면서 부분적으로 주파수 성분을 추출하는 방법이다. 주파수 분석은 시간 도메인 신호를 주파수 도메인으로 변환하여, 특정 시간대에서 주파수 성분을 분석할 수 있게 해준다. 이를 통해 차량 내 소음의 특징을 더 정밀하게 추출할 수 있다. 특히 차량 소음 분석에서는, 다양한 시간대에 발생하는 주파수 변화를 탐지하여 각 소음의 특징을 정밀하게 분석하는 데 사용된다.

STFT는 일정한 윈도우 함수 $w(t-p)$ 로 신호를 분할하고, 각 윈도우에서 신호 $x[p]$ 의 주파수 성분을 분석한다. 수식 (1)에서 제시한 STFT는 복소수 주파수 성분을 계산하는 방법으로, 이는 차량 내부의 다양한 소음 패턴을 시간-주파수 영역에서 효율적으로 분석할 수 있

도록 한다.

$$S_k(e^{j\theta}) = \sum_{p=-\infty}^{\infty} v[n-p]y[p]e^{-j\omega m} \quad (1)$$

슬라이딩 윈도우 알고리즘은 연속적인 데이터에서 중첩된 윈도우를 사용하여 데이터의 통계를 연산한다. 각 윈도우 내에서 주파수 특성을 분석한 후, 윈도우를 이동시키며 전체 신호를 분석한다. 수식 (2)에서는 윈도우 내에서 함수 $f(k)$ 를 적용하여 i 번째 윈도우의 결과를 계산한다.

$$S(i) = \sum_{k=i}^{i+N-1} f(k) \quad (2)$$

3.2 확산 모델 기반 신호 증강

확산 모델 기반 신호 증강은 초기 잡음 상태에서 목표 데이터 분포에 도달하는 확률적 과정을 역으로 추적하며, 시간에 따라 점진적으로 노이즈가 첨가된 데이터

로부터 원 데이터를 복구한다. 이는 마르코프 체인(Markov chain)에 기반하여 점진적으로 노이즈가 추가된 상태에서 데이터를 생성하고, 그 과정을 역으로 추적하여 원래의 신호를 복구하는 것이다. 이 과정에서 가우시안 노이즈가 데이터에 추가되며, 노이즈가 첨가된 데이터를 생성하고, 해당 데이터를 기반으로 한 스텝 이전 상태를 복구한다. 이 방법은 다양한 소음 데이터의 복잡한 패턴을 다루는 데 적합하며, 차량 내부에서 발생하는 다양한 주파수의 소음을 복원하는 데 사용된다.

수식 (3)에서 ϵ 은 가우시안 노이즈이고, 원본 데이터에서 노이즈를 점차적으로 추가하여 u_t 를 생성한다. 수식 (4)에서 시간 t 에서 노이즈가 첨가된 u_t 에서 한 스텝 이전의 u_{t-1} 을 추정한다. 여기서 γ_t 는 노이즈 스케줄, $\omega(u_t, t)$ 는 학습된 노이즈 모델을 통해 얻어진 노이즈 예측이다.

$$u_t = \sqrt{1-\beta_t} \cdot u_{t-1} + \sqrt{\beta_t} \cdot \epsilon \quad (3)$$

$$u_{t-1} = \frac{1}{\sqrt{1-\gamma_t}} \left(u_t - \frac{\gamma_t}{\sqrt{1-\sum_{s=1}^t \gamma_s}} \omega_t(u_t, t) \right) \quad (4)$$

3.3 이미지 유사성 비교 및 전문가 지식 기반의 클러스터링

전문가 지식 기반의 클러스터링 기법은 군집화(Clustering) 작업을 통해 확산 모델이 생성한 데이터를 가까운 군집 중심에 할당하고, 해당 군집에 전문가가 정의한 라벨을 적용하는 과정이다. 이는 랜덤하게 생성된 이미지를 일일이 라벨링하지 않으면서도 전문가의 지식을 활용할 수 있다. 특히 차량 소음 분류 작업에서는 복잡하고 불규칙적인 상황에서 소음을 분류하기 위해 고도의 확산 모델을 바탕으로 대량의 데이터셋을 생성하기 때문에 전문가가 수동으로 라벨링하기 어려운 상황이다. 따라서 논문은 군집화 과정을 통해 비슷한 특성을 가진 데이터 그룹끼리 묶고, 해당 군집에 전문가가 정의한 라벨을 할당한다.

수식 (5)에서 $PL(x_i)$ 는 각 데이터 x_i 가 속하는 군집 c_j 의 레이블 $l(c_j)$ 와 일치하며, 이는 데이터가 K 에 정의된 군집들 중 하나이다. x_i 는 유클리디언 거리의 제곱 $\|x_i - \mu_c\|^2$ 으로 측정된 가장 가까운 군집 중심 μ_c 를 찾아 군집화 분석한다. 구조화된 집합을 형성하여 데이터 간의 상대적 위치와 전체의 구조를 파악하고, 전문가 지식을 기반으로 군집의 라벨을 할당하여 데이터를 증강한다.

$$PL(x_i) = l(c_j) \mid x_i \in C_j, C_j = \arg \min_{c \in K} \|x_i - \mu_c\|^2 \quad (5)$$

이 과정에서 생성된 이미지들간 유사성을 비교하기 위해 구조적 유사도 (Structural similarity index) 척도

Algorithm 1: Pseudo-labeling with DDPM

Input:

windowed_data: array[N, F]

windowed_labels: array[N]

T: int

sigma: float

Output:

pseudolabeled_data: array[N, F]

pseudolabels: array[N]

Initialization

1: Initialize ddpm_model as Model(audio_data_shape)

2: noise_schedule \leftarrow set_noise_schedule(T, sigma)

Training Loop:

3: for epoch = 1 to num_epochs do

4: minibatch \leftarrow sample(windowed_data, batch_size)

5: noisy_minibatch \leftarrow add_noise(minibatch, noise_schedule[epoch])

6: ddpm_model.train(noisy_minibatch, minibatch)

7: if is_milestone(epoch) then

8: ddpm_model.save_weights()

Model Training

9: for i = 1 to N do

10: sample \leftarrow windowed_data[i]

11: noisy_sample \leftarrow add_noise_incrementally(sample, T, sigma)

12: denoised_sample \leftarrow ddpm_model.denoise(noisy_sample)

13: pseudolabel \leftarrow predict_label(denoised_sample)

14: pseudolabeled_data[i] \leftarrow denoised_sample

15: pseudolabels[i] \leftarrow pseudo-label

Denoising and Pseudo labeling

16: combined_data \leftarrow concatenate(windowed_data, pseudolabeled_data)

17: combined_labels \leftarrow concatenate(windowed_labels, pseudolabels)

18: train_model(combined_data, combined_labels)

19: evaluate_model(test_data, test_labels)

가 활용되며, 이는 이미지의 구조적 변화를 감지하는 방식으로 설계된 지표이다. 두 이미지 간의 구조, 밝기, 대비의 유사성을 측정한다.

수식 (6)에서는 a, b 는 비교할 두 이미지이고, 각각 원본 이미지와 모델 별 생성 이미지를 나타낸다. m_a 와 m_b 는 각각 두 이미지의 평균값, s_a^2 와 s_b^2 는 각각 두 이미지의 공분산이며 두 이미지의 유사성을 모델 별로 비교한다.

$$(a, b) = \frac{(2m_a m_b + C_1)(2s_{ab} + C_2)}{(m_a^2 + m_b^2 + C_1)(s_a^2 + s_b^2 + C_2)} \quad (6)$$

4. 실험 결과

4.1 차량 내 소음 데이터셋 수집

표 2는 차량 내부에서 수집된 소음 데이터를 제시한다. 30개 센서의 배열로 얻어 샘플을 STFT 스펙트로그램으로 변환하였다. 차원은 각각 윈도우의 총 수, 윈도우의 크기, 주파수 대역 별 강도, 그리고 센서의 수이다.

4.2 차량 내 소음 데이터 생성 및 분류 성능 비교

표 3에서는 전통적인 데이터 증강 기법과 최신 딥러닝 기반 증강 방법의 분류 성능을 다양한 노이즈 레벨

표 2 통제된 환경에서 샘플링된 차량 소음별 유형 및 길이
Table 2 Types and Duration of Vehicle Noises Sampled in a Controlled Environment

Class	Noise type	Windows Sampled	Length[s]
0	Shaker noise	332	6.8
1	Center console armrest 1	110	2.4
2	Center console armrest 2	544	11.0
3	Center console armrest 3	546	11.0
4	Seat rail 1	602	12.2
5	Seat rail 2	596	12.0
6	Seat backboard	1162	23.4
7	Sun visor	1036	20.8
8	Crash pad	610	12.3
9	Passenger seat armrest	390	7.9

표 3 기존 생성 방식과 DDPM 기반 생성 비교
Table 3 Comparison Between Traditional Generation Methods and DDPM-Based Generation

Augmentation Model	Noise level		
	0.01	0.05	0.1
Flip + CNN	0.9888	0.9751	0.9672
Stretch x 0.9 + CNN	0.9860	0.9788	0.9732
Stretch x 1.0 + CNN	0.9880	0.9830	0.9775
Stretch x 1.1 + CNN	0.9869	0.9808	0.9756
Speed x 0.9 + CNN	0.9842	0.9744	0.9686
Speed x 1.0 + CNN	0.9892	0.9823	0.9790
Speed x 1.1 + CNN	0.9878	0.9787	0.9716
GAN + CNN	0.9911	0.9838	0.9778
DDPM augmented CNN	0.9960	0.9921	0.9838

에서 비교하였다. 제안된 DDPM-augmented CNN 방식은 모든 노이즈 조건에서 가장 높은 정확도를 기록하였으며, 전통적인 증강 방식 대비 최대 1.06%p, GAN 기반 증강 방식 대비 최대 0.06%p의 성능 향상을 달성하였다.

표 4에서는 다양한 머신러닝 및 딥러닝 모델의 성능을 비교한 결과, 제안한 방법은 모든 노이즈 레벨에서 가장 높은 성능을 보였으며, CNN 기반 모델의 높은 성능 대비 최대 0.45%p의 향상이 있었다. 이를 통해 확산 모델 기반 증강 기법이 CNN의 성능을 극대화함을 알 수 있다.

그림 2에서는 데이터 생성 방법에 따른 생성된 샘플 수 변화에 따른 정확도를 비교하였다. 제안한 방법은 생성된 샘플 수가 증가할수록 정확도가 지속적으로 향상되는 경향을 보였으며, GAN을 사용한 모델과 비교할 때 최대 0.81%p의 성능 개선을 보였다. 이를 통해 제안된 방식이 GAN보다 더 효율적으로 학습 데이터를 증강할 수 있음을 확인할 수 있다.

그림 3의 SSIM 평가 결과, 제안한 확산 모델은 GAN보다 더 세부적인 특징을 잘 포착하여 더 높은 품질의 샘플을 생성하였다. 이는 본 방식이 데이터 증강에서 더 높은 품질과 성능을 제공함을 입증한다.

표 4 모델 별 분류 정확도 성능 비교
Table 4 Performance Comparison of Classification Accuracy by Model

Model	Noise level		
	0.01	0.05	0.1
KNN	0.9523	0.9492	0.9438
RF	0.9882	0.9694	0.9564
MLP	0.9882	0.9766	0.9631
LSTM	0.9561	0.9499	0.9445
Temporal Convolution	0.9882	0.9836	0.9760
CNN	0.9915	0.9885	0.9825
DDPM augmented CNN	0.9960	0.9921	0.9838

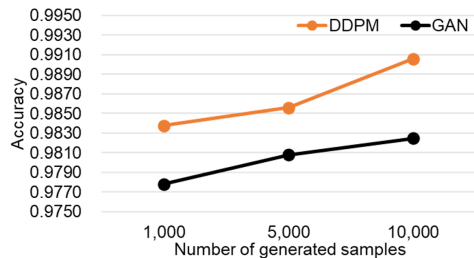


그림 2 데이터 생성 개수에 따른 모델 별 정확도
Fig. 2 Model Accuracy Based on the Number of Generated Data Points

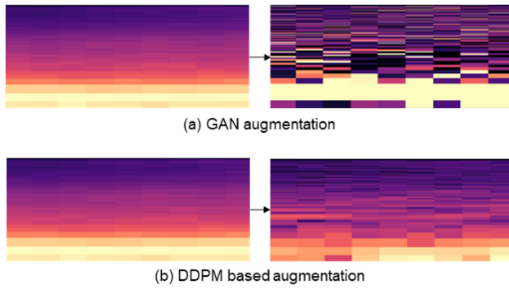


그림 3 모델 별 생성 스펙트로그램 시각화
Fig. 3 Visualization of Generated Spectrograms by Model

4.2 전문가 지식 기반의 클러스터링 적용 전후 성능 비교

그림 4에서는 논문에서 제안한 전문가 지식 기반의 군집화 과정을 적용한 전후의 증강 데이터 공간을 시각적으로 비교하였다. (a)에서는 군집화 과정이 적용되지 않은 데이터를, (b)에서는 군집화가 적용된 데이터 공간의 모습을 보여준다. 전문가 지식 기반의 군집화가 적용되지 않은 경우, 데이터의 그룹이 불명확하고 불규칙하게 분포되어 있는 반면, 군집화를 적용한 경우에는 데이터가 명확하게 구분된 군집으로 시각화된다.

이 실험은 전문가 지식 기반의 군집화를 적용함으로써 데이터 증강 과정 중 라벨링 과정의 효율성을 극대화할 수 있음을 시각적으로 증명한다. 군집화된 각 그룹은 전문가의 사전 정의된 라벨에 의해 정밀하게 분류될 수 있으며, 이를 통해 일일이 데이터를 라벨링하지 않으면서도 전문가의 지식을 활용할 수 있는 방법을 제시하였다. 특히, 소음 데이터와 같이 불규칙한 데이터 분포를 가진 환경에서 군집화된 데이터를 기반으로 효율적인 분류 성능 향상을 기대할 수 있다. 이 실험은 데이터

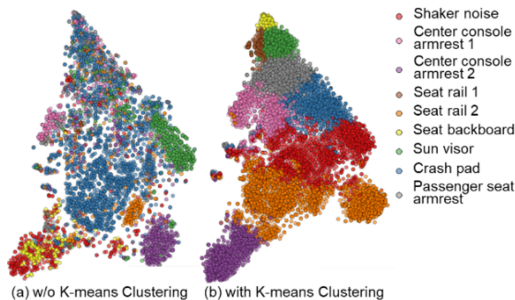


그림 4 전문가 지식 기반 클러스터링으로 라벨링된 증강 데이터 공간 시각화

Fig. 4 Visualization of an Augmented Data Space Labeled by Expert Knowledge-guided Clustering

라벨링에 소요되는 비용과 시간을 절감하면서도 성능을 향상시킬 수 있는 효과적인 방법임을 입증한다.

5. 결론 및 향후 연구

본 연구에서는 차량 내 소음 분류 성능을 향상시키기 위해 확산 모델 기반 신호 증강 방법과 전문가 지식 기반의 군집화 기법을 결합한 새로운 소음 분류 방법을 제안하였다. 확산 모델을 통해 다양한 소음 조건을 반영한 합성 데이터를 생성하고, 전문가 지식 기반의 군집화를 통해 증강된 데이터에 자동으로 라벨을 부여함으로써, 기존의 딥러닝 기반 분류 방법들과 비교해 현저한 성능 개선을 이루었다.

또한, 전문가 지식 기반의 군집화 기법을 적용함으로써 전문가의 사전 정의된 라벨을 효과적으로 반영할 수 있었으며, 일일이 데이터를 라벨링하지 않더라도 정확한 분류 성능을 유지할 수 있음을 확인하였다. 소음 데이터의 불규칙한 분포에서도 군집화된 증강 데이터를 활용한 분류 성능의 향상이 뚜렷하게 나타났으며, 이는 소음 데이터 라벨링에서 발생하는 시간적, 자원적 부담을 줄이는 데 유용한 방법임을 시사한다.

본 연구의 결과는 차량 내 소음 분류 분야에서 신호 증강과 전문가 지식 기반의 군집화 기법을 결합한 새로운 접근법이 기존 방법에 비해 더 높은 성능을 제공할 수 있음을 입증하였으며, 이를 통해 실시간 소음 모니터링 및 결함 진단 시스템에 적용할 수 있는 가능성을 보여준다.

향후 연구에서는 클래스 조건 정보를 포함한 조건부 확산 모델(Conditional Diffusion Model)을 도입하여, 특정 소음 유형을 제어 가능한 방식으로 생성함으로써 클래스 간 구분 성능을 강화하고자 한다. 또한, 능동 학습 기반의 반자동 라벨링 기법과 도메인 적용 성능 평가 지표(BWT, forgetting 등)를 추가적으로 도입함으로써, 복잡하고 변동성 높은 환경에서도 일반화 가능한 차량 소음 분류 시스템으로 확장할 수 있을 것이다.

References

[1] Cerrato, Gabriella. "Automotive sound quality--Accessories, BSR, and brakes," *Sound and Vibration*, 43:9: 10, 2009.

[2] S. J. Bu and S. B. Cho, "Classifying in-vehicle noise from multi-channel sound spectrum by deep beamforming networks," *IEEE Int. Conf. on Big Data*, pp. 3545-3552, 2019.

[3] S. J. Bu and S. B. Cho, "Automated learning of in-vehicle noise representation with triplet-loss embedded convolutional beamforming network," *Int. Conf. on Intelligent Data Engineering and Automated Learning*, pp. 507-515, 2020.

[4] K. Valev, A. Schumann, L. Sommer and J. Beyerer,

- "A systematic evaluation of recent deep learning architectures for fine-grained vehicle classification," *Pattern Recognition and Tracking XXIX*, Vol. 10649, p. 1064902, 2018.
- [5] A. Madhu and S. Kumaraswamy, "Data augmentation using generative adversarial network for environmental sound classification," *IEEE European Signal Processing Conf*, pp. 1-5, 2019.
- [6] S. J. Bu, H. J. Moon and S. B. Cho, "Adversarial signal augmentation for CNN-LSTM to classify impact noise in automobiles," *IEEE Int. Conf. on Big Data and Smart Computing*, pp. 60-64, 2021.
- [7] B. Bahmei, E. Birmingham and S. Arzanpour, "CNN-RNN and data augmentation using deep convolutional generative adversarial network for environmental sound classification," *IEEE Signal Processing Letters*, Vol. 29, pp. 682-686, 2022.
- [8] K. W. Cheng, H. M. Chow, S. Y. Li, T. W. Tsang, H. L. B. Ng, C. H. Hui, Y. H. Lee, K. W. Cheng, S. C. Cheung, C. K. Lee, S. W. Tsang, "Spectrogram-based classification on vehicles with modified loud exhausts via convolutional neural networks," *Applied Acoustics*, Vol. 205, p. 109254, 2023.
- [9] C. Tian, Y. Ma, J. Cammon, F. Fang, Y. Zhang, and M. Meng, "Dual-encoder VAE-GAN with spatiotemporal features for emotional EEG data augmentation," *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, Vol. 31, pp 2018- 2027, 2023.
- [10] Z. Zhao, H. Bai, Y. Zhu, J. Zhang, S. Xu, Y. Zhang, K. Zhang, D. Meng, R. Timofte, L. Van Gool. "DDFM: denoising diffusion model for multi-modality image fusion." *Proc. of the IEEE/CVF Int. Conf. on Computer Vision*, pp. 8082-8093, 2023.
- [11] Z. Kong, W. Ping, J. Huang, K. Zhao, and B. Catanzaro, "DiffWave: A versatile diffusion model for audio synthesis," *arXiv preprint arXiv:2009.09761*, 2020.
- [12] Y. Chen, Z. Yan, Y. Zhu, Z. Ren, J. Shen, and Y. Huang, "Data augmentation for environmental sound classification using diffusion probabilistic model with top-k selection discriminator," *Proc. Int. Conf. Intelligent Computing, Singapore: Springer Nature*, pp. 283 - 295, 2023.
- [13] J. M. Lemercier, J. Richter, S. Welker, E. Moliner, V. Välimäki, and T. Gerkmann, "Diffusion models for audio restoration: A review," *IEEE Signal Processing Magazine*, Vol. 41, No. 6, pp. 72 - 84, 2025.



백 무근

2022년~현재 경상국립대학교 컴퓨터공학부 학사과정. 관심분야는 신호처리, 딥러닝

부 석 준

정보과학회논문지
제 52 권 제 6 호 참조

최 석 훈

정보과학회논문지
제 52 권 제 6 호 참조