

딥러닝 기반 사이렌 감지 및 음원 위치 추적을 통한 자율주행차량의 긴급 차량 회피

김 찬 주

한양대학교 미래자동차공학과

Emergency Vehicle Avoidance in Self-Driving Car through Deep Learning-Based Siren Detection and Sound Source Localization

Chanju Kim

Department of Automotive Engineering, Hanyang University, Seoul 04763, Korea

Abstract : This study aims to propose a deep learning-based methodology and self-testing procedure for rapid avoidance to emergency vehicles of self-driving car. Initially, a CNN-based deep learning model was devised to detect sirens and predict the location of sound sources. To train the model, a dataset was constructed by distributing various positions of sound sources, including sirens, on a board displayed in 1cm units of 3m*3m size. Subsequently, tests were designed using an ambulance RC car with siren and an Arduino-based self-driving car. The siren detection accuracy and recall of the model were 97% and 98%, respectively. Furthermore, root mean squared error in relative positions in the x and y directions were 10.5cm and 8.5cm. In small-scale tests, fusion with vision sensors has shown effective avoidance performance for two scenarios. Therefore, deep learning-based methodology for siren detection and localization is expected to be useful in real-world self-driving systems.

Key words : Deep Learning(딥러닝), Emergency Vehicle Avoidance(긴급 차량 회피), Self-Driving Car(자율주행 차량), Self-Testing(자체 테스트), Siren Detection(사이렌 감지), Sound Source Localization(음원 위치 추적)

1. 서 론

현대의 자율주행 기술은 다양한 도로 상황에 대응하기 위해 지속적인 발전을 이루고 있습니다. 그러나 특히 긴급 차량에 대한 빠른 대응은 여전히 어려운 상황 중 하나입니다. 긴급차량이 교차로에서 건물에 가려져 보이지 않거나 후방 차량에 가려져 있는 상황 등, 현재의 비전 시스템은 특정 상황에서 한계를 드러내고 있습니다. 긴급 차량이 비전 시스템에 의해 재빠르게 감지되지 못하는 경우, 늦은 대응으로 인해 사고가 발생하거나 긴급 차량의 이동에 지연을 초래하여 피해를 발생시킬 수 있습니다.

이러한 도전에 직면하여 본 연구에서는 긴급 상황에서의 자율주행차량의 안전한 운행을 위한 혁신적인 해결책을 모색하고자 합니다. 기존의 비전 시스템만으로는 감지가 어려운 상황에서, 오디오

신호를 활용하여 긴급 차량을 효과적으로 감지하고 회피 기동을 수행할 수 있는 새로운 방법을 제안합니다.

현재의 자율주행차량이 주로 활용하는 비전 시스템은 시야에 장애물이 없을 때에만 최적의 성능을 발휘할 수 있습니다. 그러나 자율주행차량에게 시야에 제약이 있는 상황에서도 신속하게 대응할 수 있는 능력은 필수적입니다. 따라서 본 연구에서는 자율주행차량에 인간의 귀처럼 오디오 신호를 받아들일 수 있는 일련의 마이크 센서를 장착하고, 오디오 신호 속 방향과 거리로 표현되는 특성들을 학습하여 음원의 위치를 추정하는 딥러닝 모델을 개발합니다. 이를 통해 자율주행차량은 인간 운전자처럼 사이렌을 인지한 직후에 곧바로 실시간으로 추정되는 긴급 차량의 위치를 이용해 미리 회피 기동을 준비할 수 있을 것입니다.

2. 선행 연구

앞서 진행된 선행연구들에서는 마찬가지로 긴급 차량에 대해 효과적으로 대응하기 위해 다양한 시도가 있었습니다. 그 중 하나로 실시간으로 긴급 차량에 대응하기 위하여 자율주행차량의 자체적인 인지 시스템과 지능형 교통 시스템과의 통신을 모두 활용하는 커넥티드-자율주행차량 협동 시스템을 제안한 연구가 있습니다.¹⁾

해당 연구에서는 오디오 시스템을 활용하여 사이렌을 감지하고 긴급 차량의 유형을 식별하며, 시간 지연 추정 방법과 사운드 강도 프로브 방법을 사용하여 소리의 방향을 예측하는 프로세스를 제시하였습니다. 또한 이러한 오디오 기반의 정보를 비전 시스템에서 감지한 긴급 차량의 유형, 상대적인 방향 및 거리와 통합하고, 추가로 지능형 교통 시스템 및 다른 커넥티드-자율주행차량과 통신하여 최종적인 판단 및 기동을 수행하는 개념을 소개하였습니다. 그러나 해당 연구에서는 성능을 평가하기 위한 테스트가 수행되지 않은 한계가 있습니다.

또한 오디오 시스템 및 딥러닝 모델을 활용하여 사이렌을 감지하고 긴급 차량의 존재를 인지하는 일련의 두 번의 연구가 있었습니다. 첫 번째 연구에서는 원시 파형과 MFCC, 로그 멜 스펙트로그램을 입력 받아 사이렌으로의 분류 확률을 출력하는 컨볼루션 신경망의 아키텍처를 제시하고 모델의 성능을 평가하였습니다.²⁾ 더 나아가 두 번째 연구에서는 앞선 연구의 오디오 시스템 및 딥러닝 모델에 이어 긴급 차량 감지를 위한 비전 시스템 및 딥러닝 모델을 제안하였습니다.³⁾

그러나 해당 선행연구에서는 각 딥러닝 모델의 아키텍처에 따른 성능에 더욱 중점을 두고 있으며, 두 시스템 간의 유기적인 퓨전에 대한 실험결과가 부재합니다. 특히 비전 시스템과 오디오 시스템을 결합하여 실시간으로 활용하는 시스템을 제안하고 적절한 추론 빈도를 제시하는 등의 노력을 기울였지만, 두 시스템 간의 상호작용에 따른 회피 기동의 세부적인 평가가 부재한 한계가 있습니다.

주목할 만한 연구로는 앞선 연구²⁾에서 제시한 오디오 딥러닝 모델을 기반으로 사이렌 감지 및 음원 위치 추적까지 수행하는 딥러닝 모델을 제안

하고 구체적인 테스트 결과를 제공한 아폴로 사의 연구가 있습니다.⁴⁾

해당 연구에서는 직접 자율주행차량에 다수의 마이크 센서들을 부착하여 실제 도로 및 주행 환경에서 학습 데이터를 수집하였습니다. 또한 자체 자율주행 시스템에 탑재된 센티미터 수준의 정확한 측위 모듈을 사용하여 음원 위치에 대한 자동 라벨링 방법을 개발하였습니다.⁵⁾ 구체적인 성능으로는 10m~50m의 거리 범위에서 사이렌 감지 성능으로 재현율 99.16%를 달성하였습니다. 또한 음원의 상대적인 방향 추적 성능은 절대 오차의 중앙값과 평균값이 각각 9.64°, 19.18°로 나타났고, 음원의 상대적인 거리 추적 성능은 절대 오차의 중앙값과 평균값이 각각 9.30m, 10.58m로 나타나 준수한 성능을 확인할 수 있었습니다.

앞선 연구들은 자율주행차량의 긴급 차량에 대한 대응을 향상시킬 수 있는 가능성을 보여주었지만 몇 가지 한계가 있었습니다. 본 연구는 이러한 한계를 극복하기 위해 오디오 및 비전 시스템을 결합하여 자율주행차량의 실시간 긴급 차량 대응 시스템을 검증하는데 중점을 두고 있습니다. 세 가지의 시나리오에서의 자체적인 테스트를 통해 연구의 효과적인 성과를 입증하고, 다양한 상황에서의 안전한 회피 기동 능력을 확보하고자 합니다.

3. 연구 방법

이 장에서는 특히 오디오 시스템에 주목하고 있습니다. 먼저 사이렌을 감지하고 음원 위치를 추적할 수 있는 딥러닝 모델의 아키텍처를 제안합니다. 또한 이러한 모델의 학습 데이터셋을 만들기 위해 자체적으로 구축한 방법과 자율주행차량의 판단부에 필요한 룰 기반의 센서 퓨전 알고리즘도 간략하게 소개합니다. 마지막으로 아두이노 자율주행차량 및 구급차 RC카를 활용한 세 가지 시나리오에서의 회피 기동을 검증한 구체적인 테스트 과정에 대해 소개합니다.

3.1 오디오 딥러닝 모델 아키텍처

오디오 딥러닝 모델의 아키텍처는 동일한 모델을 고안한 앞선 연구⁴⁾를 참고하였습니다.

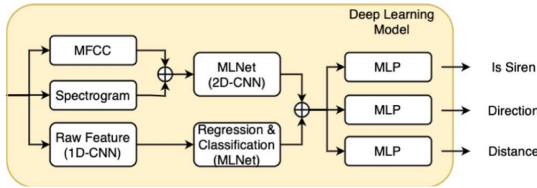


Fig. 1 Architecture of Apollo's audio deep-learning model

위의 그림에서 볼 수 있듯이 해당 연구에서 제안된 모델은 오디오 데이터의 두 종류의 특성을 처리하기 위해 두 개의 병렬 스트림을 사용했습니다. 첫 번째 스트림에서는 오디오 데이터의 주파수 도메인 특성을 학습하기 위해 MFCC와 로그 멜 스펙트로그램을 추출합니다. 두 번째 스트림에서는 원시 파형을 1D 컨볼루션 레이어에 통과시켜 시간 도메인에서의 새로운 특징 추출합니다.

이러한 두 종류의 특성은 각각 MLNet이라고 불리는 2D 컨볼루션 레이어를 통과합니다. 두 스트림의 연산 결과를 더해 3개(실제로는 4개)의 완전 연결 레이어에 통과시켜 사이렌 분류의 확률, 음원의 방향을 나타내는 \sin 값과 \cos 값, 음원의 거리를 예측하도록 설계하였습니다. 이러한 구조를 통해 모델은 주파수 도메인과 시간 도메인에서의 특성을 동시에 학습하고 이를 통합하여 준수한 예측 성능을 나타낼 수 있었습니다.

이에 본 연구는 해당 아키텍처를 참고하여 자체 구축한 데이터셋의 복잡도에 맞는 적합한 깊이를 가질 수 있도록 모델 아키텍처를 고안하였습니다. 이 과정에서는 두 가지 모델 아키텍처 유형을 고려해야했습니다. 첫 번째로 한 개의 통합된 딥러닝 모델로 사이렌 감지와 음원 위치 추적을 동시에 수행하는 이른바 단일 단계 모델이 있습니다. 이 모델의 장점은 추론 속도가 빠르다는 것입니다. 두 번째로 두 개의 연속적인 딥러닝 모델을 사용하여 첫 번째 모델에서는 사이렌 감지를 수행하고 사이렌으로 판단된 경우 두 번째 모델에서 음원 위치 추적을 수행하는 이중 단계 모델이 있습니다. 이 모델에는 두 번째 모델에서 사이렌의 거리 데이터에만 담겨있는 일관된 패턴을 학습할 수 있다는 장점이 있습니다. 추론 속도를 중요시한 본 연구에서는 먼저 단일 단계 모델을 설계하여 성능을 평가하였습니다. 그러나 자체 구축한 데이터셋의 한계로 인해 앞서 서술한 장점을 가지는 이중 단계 모델을 채택하였습니다.

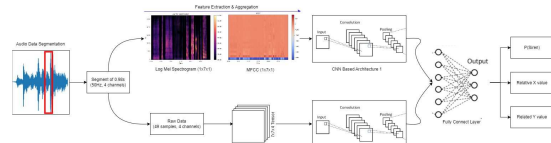


Fig. 2 Architecture of our audio deep-learning model

먼저 본 연구에서 시도해본 다양한 모델 아키텍처들은 동일한 형태의 두 종류의 입력 텐서를 사용합니다. 먼저 각기 다른 방향으로 배치된 4개의 마이크 센서에서 50Hz의 샘플링 레이트로 오디오 신호를 받아들였습니다. 49개 샘플을 가지고 있는 0.98초의 오디오 세그먼트 4개(마이크 당 하나)를 사용하였고, 이를 원시 파형을 나타내는 (7, 7, 4) 크기의 텐서와 (1, 7, 1) 크기의 MFCC와 로그 멜 스펙트로그램을 붙인 (1, 14, 1) 크기의 텐서를 사용했습니다. 이중 단계의 두 모델 모두 동일한 아키텍처를 사용하며 컨볼루션 레이어의 깊이나 필터 크기 등 세부적인 사항에는 차이가 있습니다.

3.2 마이크 센서 및 샘플링 레이트



Photo. 1~3 LM393, KY-037 and MAX9814

처음에는 LM393이나 KY-037과 같은 증폭 기능이 없는 마이크 센서 모듈을 사용했습니다. 그러나 실제 출력되는 오디오 데이터를 확인하며 소리 감지 범위가 테스트에 사용되기에는 터무니없이 부족한 약 20cm 미만에 불과하다는 것을 확인했습니다. 따라서 본 연구에서는 약 1.5m 이상의 소리 감지 범위 및 증폭 기능을 가지고 있는 MAX9814를 사용하였습니다.

MAX9814의 주파수 범위는 20Hz~20KHz이지만 아두이노 연산을 수행한 후 컴퓨터로 전송된 샘플링 레이트는 최대 50Hz에 불과하였습니다. 이에 대표적인 오디오 데이터 전처리 라이브러리인 Librosa를 활용하여 5KHz~22.5KHz의 범위 내에서 리샘플링을 수행해 데이터의 길이를 늘리고자 시도하였습니다. 그러나 제한된 CPU 및 GPU 메모리로 인해 데이터를 전처리하고 모델을 학습하는데 과부하가 걸리는 현상이 발생하여 중단하였고, 원래대로 50Hz를 사용하여 모델을 학습시켰습니다. 또한 MAX9814의 분해능은 10bit로

사운드 강도를 1024단계로 구분된 전압으로 표현합니다. 본 연구에서 또한 dB 단위가 아닌 0~1024의 범위로 사운드 강도를 표현하였습니다.

3.3 오디오 데이터셋 구축

본 연구에서 필요로 하는 사이렌의 음원 위치가 라벨링 되어있는 데이터셋을 찾을 수 없어 직접 제작하였습니다. 약 20cm~30cm의 전장을 가진 아두이노 자율주행차량과 구급차 RC카를 활용하여 수행할 테스트를 대비해, 3m*3m 크기의 적절한 테스트 범위를 선택했습니다. 이는 테스트 차량의 약 10배의 크기로, 약 5m의 전장을 가지는 일반 차량의 실제 크기 기준으로 약 50m*50m의 범위에서 테스트하는 것과 동일하며 앞선 연구⁴⁾에서 사용한 10~50m의 거리 범위에 부합합니다.

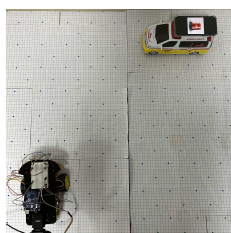


Photo. 4 Dataset constructing environment with self-created board

아두이노 자율주행차량과 음원 사이의 상대적인 위치를 정확히 측정하기 위해, 1cm 간격으로 표시된 보드를 수작업으로 제작하였습니다. 학습 데이터셋을 실제로 구축할 때 라벨링된 음원 위치는 아두이노 자율주행차량과 구급차 RC카의 전면 중앙부, 후면 중앙부 중 서로 가까운 지점을 기준으로 측정되었습니다. 이러한 방식은 실제 음원 위치를 벗어나는 차량의 크기로 인한 충돌 가능성을 줄이기 위해 보수적으로 설정한 것이었으나, 실제로는 구급차가 다가오는 상황과 멀어지는 상황에서 라벨링 기준이 일관되지 못한 문제점을 발견하였습니다. 이러한 문제가 오디오 딥러닝 모델의 성능에 부정적인 영향을 끼쳤다고 판단됩니다.

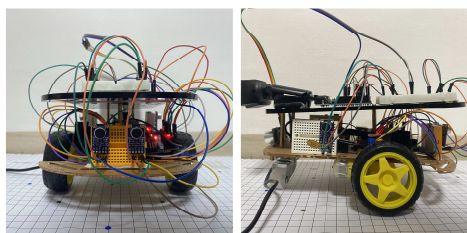


Photo. 5~6 Arduino-based self-driving car

위의 사진에서 확인할 수 있듯이, 자체 제작한 아두이노 자율주행차량은 총 4개의 마이크 센서를 장착하고 있습니다. 마이크 센서는 전방에 대칭으로 2개, 양 측면에 대칭으로 2개가 위치하고 있습니다. 각각의 마이크 센서들은 장착된 위치와 음원의 위치에 따라 각기 다른 사운드 강도로 오디오 신호를 받아들이게 됩니다. 간단한 소리를 발생시켜 확인해 본 결과, 전후방보다 좌우 방향의 소리를 구분하는 것에 어려움을 가지는 경향이 있었습니다. 이에 따라 후방에 마이크를 배치하는 대신 전방에 좌우 방향으로 마이크를 배치하여 구분 성능을 향상시키고자 하였습니다.

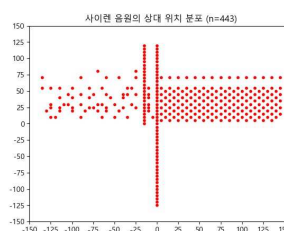


Fig. 3 Relative position distribution of siren sound source

먼저 사이렌 음원으로는 유튜브에 올라와 있는 ‘Ambulance siren sound - 1 Hour’ 영상⁶⁾을 사용하였습니다. 위의 그림에서 확인할 수 있듯이, 음원은 아두이노 자율주행차량을 기준으로 전후좌우 모두 1.5m 거리 이내에 위치시켰습니다. 음원의 위치는 x, y축 각각 -150cm~150cm의 범위 내에서 라벨링하였습니다. 시간적인 제약으로 인해 3가지 시나리오에서 필요로 하는 음원 위치를 위주로 데이터셋을 구축했습니다. 주로 교차로 상황을 나타내는 전방 좌우 방향과 후방 동일차선 상황 등을 나타내는 전후 방향에 음원 위치가 집중돼있고, 증강시키기 전 데이터의 개수는 443개입니다.

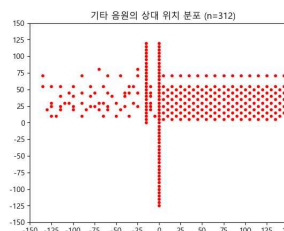


Fig. 4 Relative position distribution of other sound source

마찬가지로 사이렌이 아닌 기타 음원 또한 유튜브의 영상을 활용했습니다. 기타 음원에는 노래나 사람의 대화 소리, 레이싱 주행 소리 등 다양한 종류의 음원이 포함되었습니다. 또한 사이렌 음원의

위치 분포와 비슷하게 데이터셋을 구축했습니다. 기타 음원의 증강시키기 전 데이터의 개수는 312개입니다.

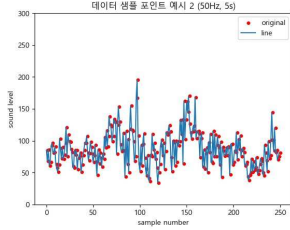


Fig. 5 Example of audio data

마이크 센서를 통해 각 음원 위치마다 5초 길이의 데이터를 수집하였습니다. 5초 동안 총 250개의 샘플 포인트가 생성되었고, 48개의 샘플 포인트를 오버랩하여 음원 위치 당 202개의 49개의 샘플 포인트를 가지는 오디오 세그먼트 데이터를 얻을 수 있었습니다. 그 결과 사이렌 음원과 기타 음원의 데이터가 각각 89,486개, 63,024개로 증강되었으며, 총 개수가 152,510개로 오디오 딥러닝 모델을 학습시키기 위해 충분하였습니다. 사이렌 음원과 기타 음원의 비율은 각각 59%, 41%로 균형 잡힌 이진 분류 데이터셋을 구축하였습니다.

3.4 시나리오 선정

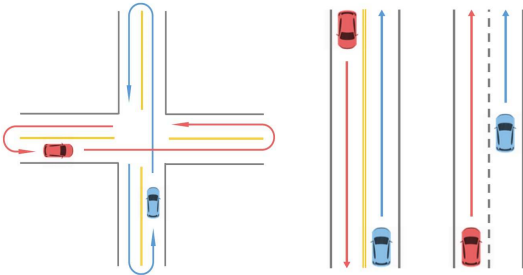


Fig. 6 Three test driving scenarios

딥러닝 기반의 오디오 및 비전 시스템을 결합한 긴급 차량 회피 시스템을 평가하기 위해 3가지 시나리오를 선정하였습니다. 위의 그림에서 아두이노 자율주행차량은 파란색 차량으로, 긴급차 RC카는 빨간색 차량으로 표시되었습니다.

첫 번째 시나리오는 긴급 차량이 교차로의 좌측 90° 방향에서 직진 주행하며 접근하는 상황입니다. 두 차량 사이의 교차로 모서리에는 건물이 세워져 있어 비전 시스템으로 긴급 차량을 인지할 수 없다고 가정합니다. 이 때 오디오 시스템을 통해 빠르게 사이렌 감지 및 음원 위치 추적을 하는 것이 중요합니다. 긴급 차량이 일정 범위 이내에

접근했다고 판단하면 감속 및 정지를 수행합니다.

두 번째 시나리오는 긴급 차량이 후방의 동일한 차선에서 접근하는 상황입니다. 위의 그림에서는 3번째에 표시되었으며, 그림과 달리 동일한 차선의 상황으로 선정하였습니다. 자율주행차량과 긴급 차량 사이에 다른 차량이 있어 비전 시스템으로 긴급 차량을 인지할 수 없다고 가정합니다. 마찬가지로 오디오 시스템을 통해 긴급 차량을 빠르게 인지하여 일정 범위 이내에 접근했다고 판단하면 차선 변경을 수행합니다.

세 번째 시나리오는 긴급 차량이 반대 차로에서 전방으로 접근하는 상황입니다. 이 시나리오에서 자율주행차량은 회피 기동을 하지 않고 정속 주행해야 합니다. 긴급 차량의 동일 차선 여부를 오디오 시스템을 통해 판단하여 두 번째 시나리오와 구분할 수 있는지 확인하기 위해 선정하였습니다.

3.5 판단 알고리즘

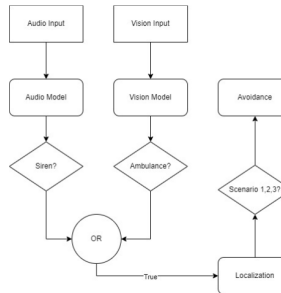


Fig. 7 Diagram of decision algorithm

다음과 같은 간단한 룰 기반의 알고리즘을 구현하여 오디오 및 비전 시스템의 판단 알고리즘으로 사용하였습니다. 먼저 오디오 및 비전 센서를 통해 데이터를 받아들인 후, 오디오와 비전 모델 각각에서 긴급차의 여부를 판단하고 OR 게이트를 통과시킵니다. 이에 따라 한쪽에서만 긴급차의 존재를 감지해도 위치 추적 단계가 수행됩니다.

위치 추적 단계에서는 오디오 시스템의 경우 기존에 입력값으로 받았던 데이터를 두 번째 모델에 다시 입력하게 되고, 비전 시스템의 경우 초음파 센서를 이용해 거리를 측정합니다. 이후 두 개의 위치 정보를 적절히 혼합해 시나리오 1, 2, 3에 해당하는지 여부를 확인하고, 시나리오가 확인되면 아두이노에 신호를 보내 적절한 제어 알고리즘을 실행시킵니다. 이러한 알고리즘은 약 1~2초 주기로 반복적으로 수행되며 딥러닝 모델의 오차로 인한 불안전성을 실시간성을 통해 보완합니다.

여기에 사용된 비전 모델은 ImageNet 데이터

셋으로 사전 학습된 ResNet50 분류 모델을 파인 튜닝한 모델입니다. 백본 모델에 추가적인 레이어를 붙인 후 약 22,000개의 자체적으로 구축한 구급차 이진 분류 데이터셋으로 학습 및 평가를 수행했습니다. 카메라 센서로는 웹캠을 사용하였고 초음파 센서 또한 부착하였습니다.

4. 연구 결과

4.1 마이크 센서 별로 측정한 음원 위치에 따른 사운드 강도

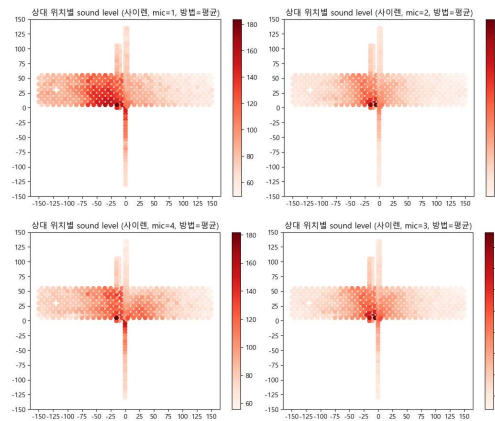


Fig. 8~11 Measured sound level by each microphone sensors based on sound source location

위의 그림들은 마이크 센서 별로 측정한 음원 위치에 따른 사운드 강도를 나타냅니다. 아두이노 자율주행차량의 좌측, 전방 좌측, 전방 우측, 우측에 부착된 각 마이크 센서가 각각 1, 2, 3, 4번 마이크를 나타냅니다. 각 마이크 센서의 사운드 수신 감도가 정확히 일치하지 않기 때문에 음원 위치에 따라서 가장 가까운 마이크 센서에서 가장 큰 음원 강도를 측정하는 패턴을 확인할 수 없었습니다.

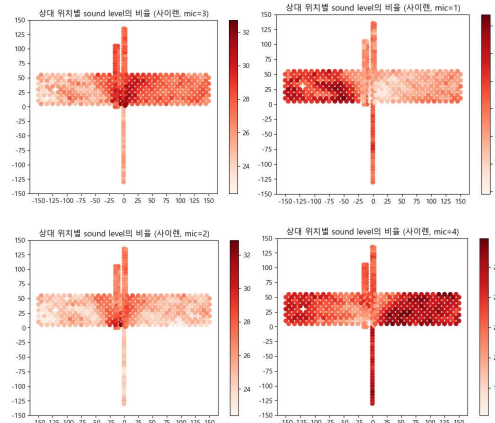


Fig. 12~15 Measured sound level ratio by each microphone sensors based on sound source location

이에 따라 각각의 음원 위치 별로 모든 마이크 센서가 받아들인 음원 강도의 총합에 대한 각 마이크 센서가 받아들인 음원 강도의 비율을 나타내었습니다. 그 결과 음원 위치에 따라서 가장 가까운 마이크 센서에서 가장 큰 음원 강도를 측정하는 패턴을 확인할 수 있었습니다.

4.2 오디오 딥러닝 모델 테스트 성능

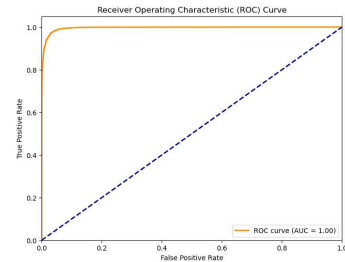


Fig. 16 ROC curve of audio classification deep-learning model

먼저 이진 분류를 통해 사이렌 감지를 수행하는 첫 번째 딥러닝 모델의 ROC 커브는 다음과 같습니다. AUC가 1.00으로 매우 좋은 성능을 보였으며, 이외의 지표에서도 정확도 0.97, 정밀도 0.97, 재현율 0.98, F1 Score 0.98로 좋은 성능을 나타냈습니다.

또한 사이렌의 음원 위치 추적을 수행하는 두 번째 딥러닝 모델의 x축 방향, y축 방향의 rmse는 각각 10.5cm와 8.5cm로 준수한 성능을 나타내었습니다. 다만 이는 음원이 정지한 상태로 수집된 데이터로 만들어진 테스트셋에서 평가된 성능이기 때문에, 실제 구급차 RC카 및 음원을 움직이며 실시간으로 음원 위치를 추적할 때에는 강건한 성능을 나타내기 어려웠습니다.

4.3 시나리오 기반 회피 기동 결과

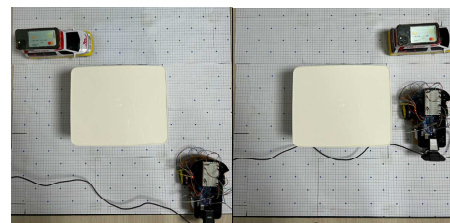


Photo. 7~8 Test setting and result of scenario 1

시나리오 1번에 대해서 테스트한 결과, 아두이노 자율주행차량이 구급차 RC카가 건물에 가려져 보이

지 않는 상황에서도 오디오 시스템을 통해 존재와 위치를 인지할 수 있었습니다. 구급차 RC카가 특정 범위 이내에 존재한다고 판단한 직후, 아두이노 자율주행차량은 성공적으로 감속 및 정지하여 구급차 RC카에 대해 적절히 회피하였습니다.

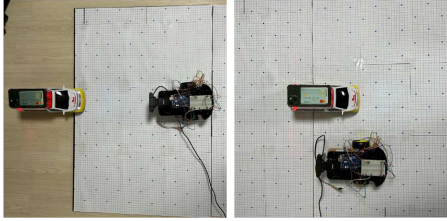


Photo. 9~10 Test setting and result of scenario 2

시나리오 2번에 대해서 테스트한 결과, 아두이노 자율주행차량이 오디오 시스템을 통해 구급차 RC카가 후방 동일 차선에서 접근한다고 판단할 수 있었습니다. 구급차 RC카가 특정 범위 이내에 존재한다고 판단한 직후, 아두이노 자율주행차량은 성공적으로 우측 차선으로 주행 경로를 변경하여 구급차 RC카에 대해 적절히 회피하였습니다.

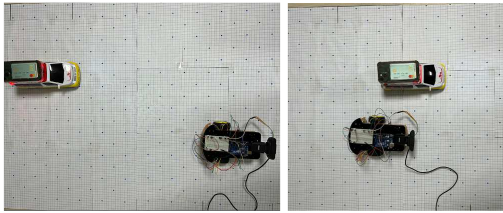


Photo. 11~12 Test setting and result of scenario 3

시나리오 3번에 대해서 테스트한 결과, 아두이노 자율주행차량이 오디오 시스템을 통해 구급차 RC카가 전방 반대 차로에서 접근한다고 판단할 수 있었습니다. 구급차 RC카가 직진 경로에 있지 않다고 판단한 직후, 아두이노 자율주행차량은 성공적으로 정속 주행을 이어갔습니다.

5. 결 론

5.1 연구 결론

- 1) 여러 개의 마이크 센서를 각기 다른 방향으로 배치하여 음원의 방향 정보가 담겨있는 학습 데이터를 구축할 수 있었습니다.
- 2) 음원 위치에 따른 사이렌 음원에 대한 사운드 강도를 통해 음원의 거리가 담겨있는 학습 데이터를 구축할 수 있었습니다.
- 3) 위와 같은 학습 데이터셋을 이용해 학습시킨 딥러닝 모델을 통해 음원의 위치를 추정할 수 있었습니다.

4) 위와 같은 오디오 시스템과 비전 시스템을 아두이노 자율주행차량에 적용하여 다양한 시나리오에 대해 성공적으로 회피 기동을 수행하는 것을 확인할 수 있었습니다.

5) 자체적으로 학습 데이터 및 테스트 환경을 구축하여 딥러닝 모델 및 판단 알고리즘을 평가하는 방법을 제안하였습니다.

5.2 한계

- 1) 사운드 감지 성능이 좋지 않은 마이크를 사용하여 약 75cm 이상의 거리의 음원에 대해서는 측정된 사운드 강도의 차이가 크지 않았습니다.
- 2) 실제로 음원이 이동하는 점을 데이터에 반영시키지 못하였음. 구급차 RC카의 속도에 따라 추적된 위치와 실제 위치와의 오차가 발생하였습니다.
- 3) 샘플링 레이트의 한계에 따라 데이터에 풍부한 특성이 담기지 못하였습니다.
- 4) 음원의 위치를 라벨링 하는 방법에서 일관되지 못하다는 문제점이 발생하였습니다.
- 5) 다양한 노이즈 비율을 조절하여 강건한 데이터셋을 구축해보지 못하였습니다.
- 6) 오버랩 비율이 약 98%(48/49)로 모델이 과적합 될 수 있습니다.

5.3 기대효과

저성능 및 적은 개수의 마이크 센서와 충분하지 못한 학습 데이터셋을 사용했음에도 성공적으로 회피 기동을 수행했습니다. 이러한 점에서 실제 자율주행차량에 고성능의 마이크 센서와 충분한 데이터셋을 사용한다면 충분히 긴급 차량에 대한 안정적인 회피 기동으로 이어질 수 있다고 기대합니다.

최근 엔비디아가 주도하고 있는 GM, 닛산, 다임러 등이 포함된 컨소시엄에서도 마이크 센서를 활용하여 긴급 차량의 회피 기동을 수행하는 것에 대해 연구를 진행하고 특허를 출원했으며, 실제 자율주행차량에 마이크 센서를 부착하는 행보를 보이고 있습니다. 해당 특허에서는 간단한 컨볼루션 레이어를 통해 사이렌을 감지하고 시간 지연 추정 방법을 통해 음원 위치를 추적하는 것을 기술하고 있습니다.⁷⁾ 이에 시간 지연 추정 방법보다 더욱 빠르고 강건한 딥러닝 모델을 사용한다면 보다 안정적인 회피 기동을 할 수 있을 것으로 기대합니다.

또한 공사장 소음, 사람의 대화 소리 등 주행 상황에 인지할 수 있는 다양한 음원에 대해서 분류 및 위치 추적을 통해 다양한 객체를 인지할 수 있을 것으

로 기대합니다. 이러한 시스템에 필요한 것은 다수의 마이크 센서 및 연산 장치, 통신 장치뿐이기 때문에 자율주행차량이 아닌 일반 차량에게도 좋은 기회가 될 것으로 기대합니다.

후 기

먼저 약 11개월의 기간 동안 함께 딥러닝에 대해 공부하고 공동 연구를 진행한 김동욱 학우님의 노고에 감사드립니다. 또한 테스트 환경에 필요한 장소 제공에 협조한 김호준님, 아두이노 부품을 마련하는데 협조해주신 동신 전자 사장님께도 감사 인사를 드립니다. 이 연구가 추후 완전한 자율주행 시대가 도래하는데 기여하길 기원하며 본 연구를 마무리합니다. 참고로 본 연구에서 사용한 비전 모델에 대해서는 김동욱 학우님의 ‘구급차 회피 기동을 위한 딥러닝 및 컴퓨터 비전 기반 센서 퓨전 알고리즘’ 연구 논문에서 확인할 수 있습니다. 또한 시나리오 1~2번에 대한 테스트 주행 영상 2개를 ‘@ing903’ 아이디의 유튜브 쇼츠에서 확인할 수 있습니다.

References

- 1) Y. H. Liu, O. P. Albuquerque, P. C. K. Hung, H. A. Gabbar, M. Fantinato and F. Iqbal, “Towards a Real-Time Emergency Response Model for Connected and Autonomous Vehicles,” CIKM 2022 Workshops co-located with 31st ACM International Conference, 2022.
- 2) V. T. Tran and W. H. Tsai, “Acoustic-Based Emergency Vehicle Detection Using Convolutional Neural Networks”, IEEE Access, Vol.8, pp.75702–75713, 2020.
- 3) V. T. Tran and W. H. Tsai, “Audio-Vision Emergency Vehicle Detection”, IEEE Sensors Journal, Vol.21, No.24, pp.27905–27917, 2021.
- 4) H. Y. Sun, X. Y. Liu, K. C. Xu, J. H. Miao and Q. Luo, “Emergency Vehicles Audio Detection and Localization in Autonomous Driving”, <https://arxiv.org/abs/2109.14797>, 2021.
- 5) Apollo-Platform, ApolloAuto / apollo, <https://github.com/apolloauto/apollo>, 2023.
- 6) Filmcrew Sound Effects, “Ambulance siren sound - 1 Hour”, https://www.youtube.com/watch?v=xtfAX_6ltUA, 2023.
- 7) NVIDIA Corporation, “Emergency Respons

e Vehicle Detection for Autonomous Driving Applications”, <https://image-ppubs.uspto.gov/dirsearch-public/print/downloadPdf/20220157165>, 2020.