

클래스 불균형 환경에서 2D, 1D 멀티모달 센서 퓨전 연구

윤인영¹⁾, 김건훈¹⁾, 백재순¹⁾, 최준원²⁾
한양대학교 전기공학과¹⁾, 서울대학교 전기공학과²⁾

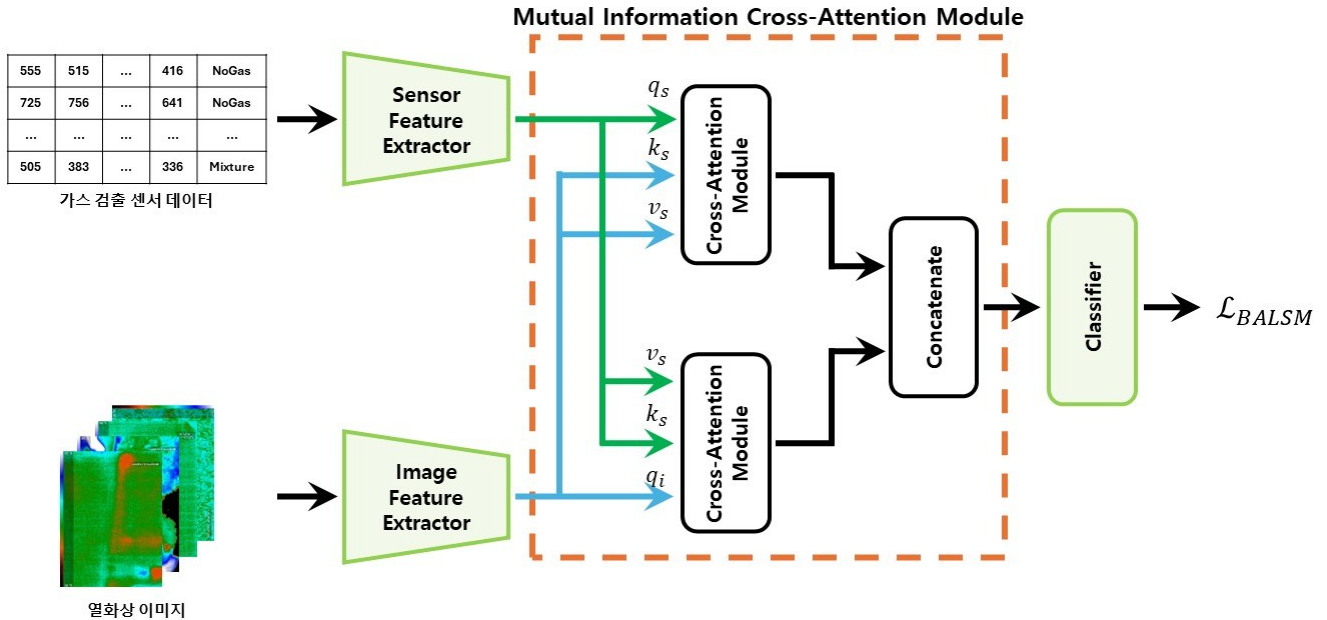
inyoungyoon@spa.hanyang.ac.kr, khkim@spa.hanyang.ac.kr, jsbaik@spa.hanyang.ac.kr,
*junwchoi@snu.ac.kr

Multimodal Sensor Fusion of 2D and 1D Data for Class Imbalance

In Young Yoon, Kun Hoon Kim, Jae Soon Baik, Jun Won Choi*
Hanyang Univ., *Seoul Univ.

요약

본 논문은 클래스 불균형 환경에서 이미지 데이터와 수치형 데이터에 담긴 정보를 융합하여 정확하고 강인한 상태 추정 모델을 제안한다. 제안된 상호 정보 교차 집중 모듈은 두 모달리티의 정보를 효과적으로 융합하기 위하여 교차 집중 모듈 두 개와 평행 융합으로 구성되어 있다. 각 모달리티의 특징 벡터가 제안된 모듈에 입력되면 상호간 유의미한 정보를 강조한 퓨전 특징 벡터를 출력한다. 또한 클래스 불균형 문제를 완화하기 위해서 균형 소프트맥스 함수를 이용하였다. 실험은 다중 센서 가스 검출 데이터셋에 대해서 수행되었으며, 단일 모달리티를 사용한 모델에 비해 높은 성능을 달성하였다.



[그림 1] 제안하는 네트워크 전체 구조

I. 서론

최근 자율주행, 제조 공정, 바이오 등 다양한 분야에서 센서 정보는 시스템의 상태를 추정하고 모니터링하는 데 활용되고 있다. 이에 따라 다양한 센서로부터 획득한 이질적인 데이터를 효과적으로 융합하기 위한 멀티모달 학습 기반의 센서 퓨전 기법들이 활발히 연구되어 왔다. 그러나 기존의 센서 퓨전 연구들은 주로 자율주행 환경에서 수집된 데이터를 대상으로 수행되었다[1]. 특히 이미지와 같은 비정형 데이터와 수치형 데이터를 융합하는 방법론에 대한 연구는 상대적으로 연구되지 않았다.

한편, 많은 인공지능 응용 분야에서 모델 학습 시 클래스 불균형(Class imbalance)은 일반적으로 나타나는

문제이다. 클래스 불균형이란 가지고 있는 샘플의 수가 극단적으로 많은 소수의 헤드 클래스(head classes)와 가지고 있는 샘플의 수가 극단적으로 적은 다수의 테일 클래스(tail classes)로 구성된 데이터 환경이다. 이러한 클래스 불균형 환경에서 학습된 모델은 헤드 클래스에 대해 편향되어 테일 클래스에 대한 예측 성능이 저하되므로, 전체적인 성능 하락을 겪게 된다. 그러나 기존의 센서 퓨전 연구들은 이와 같은 클래스 불균형 문제를 간과한 채 진행되어 왔다.

본 논문에서는 클래스 불균형 환경에서의 2D 이미지 데이터와 1D 수치형 데이터의 멀티모달 센서 퓨전 기법을 제안한다. 이미지 데이터와 수치형 데이터는 본질적으로 상이한 형태로 수집되므로, 특징 벡터 간의 융합을 위해 후기 융합(Late fusion) 방식을 채택하였다.

또한, 각 데이터 모달리티로부터 추출된 정보를 상호 보완적으로 활용하기 위해 교차 집중(Cross-attention) 모듈[4]을 도입하였다. 이에 더해 클래스 불균형 문제를 해소하기 위해 균형 소프트맥스(Balanced Softmax) 손실 함수[2]를 사용하여 모델을 학습하였다. 제안된 방법론을 통해 단일 모달리티 데이터만을 사용하는 기존 모델 대비 유의미한 성능 향상을 달성하였다.

II. 본론

2.1) 상호 정보 교차 집중 모듈(Mutual Information Cross-Attention Module)

제안하는 네트워크는 이미지 정보와 수치형 데이터의 정보를 의미있게 활용하도록 구성하였으며, 전체 구조는 [그림 1]과 같다.

이미지 데이터와 수치형 데이터에 내재된 정보를 각각 학습하기 위해 별도의 특징 벡터 추출기를 사용하였다. 추출된 두 특징 벡터는 평행 융합(concatenation) 방식을 통해 퓨전 특징 벡터로 결합된다. 이후 융합된 퓨전 특징 벡터는 분류기에 입력되어 추정 상태를 출력한다. 이 과정에서 상호 정보 교차 집중 모듈을 통해 두 특징 벡터 간의 유의미한 정보가 강조된다. 제안된 모듈은 각 모달리티에 대한 두 개의 교차 집중 모듈과 평행 융합으로 구성되어 있다. 이미지 특징 벡터가 쿼리로 입력되는 교차 집중 모듈에는 수치형 데이터 특징 벡터가 키와 값으로 입력된다. 수치형 데이터에 대해서도 동일한 과정이 적용된다. 최종적으로 두 교차 집중 모듈에서 출력된 특징 벡터는 평행 융합을 통해 퓨전 특징 벡터로 결합된다.

2.2) 균형 소프트맥스 손실 함수(Balanced Softmax Loss Function)

클래스 불균형 문제는 대다수의 샘플을 차지하는 헤드 클래스에 대한 성능은 유지되고, 상대적으로 적은 수의 샘플을 가지는 클래스의 성능은 저하되는 현상이다. 본 논문에서는 클래스 불균형 문제를 완화하기 위해 균형 소프트맥스 손실 함수[2]를 사용하였다. 기존의 교차 엔트로피 손실 함수는 클래스 불균형 환경에서 모델이 헤드 클래스로 편향되도록 학습시키는 문제가 있다. 균형 소프트맥스 손실 함수는 교차 엔트로피 손실 함수에 균형 소프트 맥스 함수를 입력한 형태로 정의된다. 구체적으로, 균형 소프트맥스 함수는 소프트맥스 함수의 분모와 분자의 지수함수에 해당 클래스의 샘플 수를 곱해준 형태로 정의된다. 이 방법은 적은 계산량 증가로 클래스 불균형 문제를 효과적으로 완화할 수 있다.

2.3) 실험 환경 및 실험 결과

2.3.1) 실험 환경

제안하는 모델의 학습 및 성능 평가를 위해 다중 센서 가스 검출 데이터셋을 활용하였다. 해당 데이터셋은 가스 누출 감지 및 누출된 가스의 종류를 분류하기 위해 수집된 것으로, 총 4 가지 클래스(NoGas, Gas Type 1, Gas Type 2, Mixture)로 구성되어 있다. 각 클래스 당 1,600 개의 샘플이 포함되어 있어, 전체 데이터셋은 6,400 개의 샘플로 이루어져 있다. 각 샘플은 열화상 카메라를 통해 획득한 열화상 이미지와 7 개의 가스 검출 센서로부터 얻은 7 개의 센서 값으로 구성된다.

[표 1] 다중 센서 가스 검출 데이터셋에 대한 실험 결과

Method	Accuracy (%)
Sensor	67.81
Vision	79.38
Our Method	93.05

본 논문에서는 클래스 불균형 환경을 모사하기 위해 NoGas 클래스를 제외한 나머지 세 클래스(Gas Type 1, Gas Type 2, Mixture)의 샘플 수를 NoGas 클래스 샘플 수의 10%로 조정하였다. 이를 통해 실제 응용 환경에서 흔히 발생하는 클래스 불균형 문제를 재현하고, 클래스 불균형 환경에서 제안하는 모델의 성능을 평가하고자 하였다.

수치형 데이터를 위한 특징 벡터 추출기로 5 레이어 DNN 을 이용했고, 이미지 데이터를 위한 특징 벡터 추출기로 MobileNet V2[3]를 이용했다.

2.3.2) 실험 결과

실험 결과는 클래스 분류 결과의 정확도를 성능 지표로 한다. [표 1]에서 Sensor 는 수치형 데이터만 이용한 모델의 성능이고, Vision 은 이미지 데이터만 이용한 모델의 성능이다. 제안된 모델은 수치형 데이터만 이용한 모델에 비해 25.24% 만큼, 이미지 데이터만 이용한 모델에 비해 13.67% 만큼 성능 향상을 달성했다.

III. 결론

본 논문에서는 클래스 불균형 환경에서도 정확하고 강인한 다중 센서 융합 모델을 제안한다. 제안하는 네트워크는 상호 정보 교차 집중 모듈로 서로 다른 모달리티에서 얻은 정보 중 유의미한 정보를 강조하고, 이를 평행 융합한다. 제안된 모델은 다중 센서 가스 검출 데이터셋에서 학습 및 성능 평가가 이루어졌으며, 단일 모달리티만 이용하는 모델 대비 매우 높은 성능 향상을 이루었다.

ACKNOWLEDGMENT

"이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2020R1A2C2012146)"

참 고 문 헌

- [1] Huang, Keli, et al. "Multi-modal sensor fusion for auto driving perception: A survey." *arXiv preprint arXiv:2202.02703* (2022).
- [2] Ren, Jiawei, et al. "Balanced meta-softmax for long-tailed visual recognition." *Advances in neural information processing systems* 33 (2020): 4175-4186.
- [3] Sandler, Mark, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
- [4] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).