

지각의 새로운 지평을 연 멀티모달 AI

김말희 한국전자통신연구원 환경ICT연구실 책임연구원
허태욱 한국전자통신연구원 환경ICT연구실 실장
이일우 한국전자통신연구원 산업에너지융합연구본부 본부장

1. 머리말

19세기 말, 인상파 화가 폴 세잔(Paul Cézanne)은 사물의 본질을 탐구하기 위해 모든 형태를 단순한 기하학적 형태로 환원시키는 접근법을 사용했다. 이는 사물의 핵심적인 특징을 추출해 표현하려는 시도이며, 이후 파블로 피카소(Pablo Picasso)의 입체주의로 확장됐다. 이러한 본질 추구 노력은 오늘날 멀티모달 AI(Multimodal AI) 발전과 유사한 점이 있다. 멀티모달 AI는 다양한 데이터 유형을 공통된 표현 공간으로 투사시켜, 복잡한 데이터를 단순화하고 핵심적인 특징을 추출해 분석한다. 이는 세잔이 기하학적 형태로 사물을 단순화함으로써 본질을 표현하려는 예술적 접근과 일맥상통하며, 서로 다른 데이터 간의 의미와 관계를 더 깊이 이해하는 데 기여하고 있다.

멀티모달 AI는 텍스트, 오디오, 이미지, 비디오 등 다양한 데이터 유형을 통합해 처리하고 분석할 수 있는 AI 기술을 의미한다[1]. 이러한 기술은 분석력 강화, 사용자와의 풍부한 인터페이스 제공이라는 특징을 갖는다. 멀티모달 AI는 기존 단일 양식 기반 AI보다 더 풍부하고 정확한 이해와 예측을 가능케 하며, 복잡한 실제 문제를 해결하는 데 중요한 역할을 한다.

또한, 멀티모달 AI는 텍스트 기반 입력으로 이미지 기반 출력을 생성하거나, 이미지 데이터와 오디오 데이터를 결합해 새로운 이미지 데이터를 생성하는 등 입력과 출력의 데이터 양식이 다양하다. 덕분에 사용자와 다양한 형태의 상호작용이 가능하며, 이를 통해 사용자 경험을 한층 강화시킬 수 있다. 이러한 이유로 멀티모달 AI는 대화형 AI(Conversational AI)의 핵심기능으로 주목받고 있다.

최근 멀티모달 AI는 자연어 처리를 비롯한 다양한 콘텐츠 생성 분야에서 큰 주목을 받고 있으며, 의료, 자율주행, 콘텐츠 등 산업 전반에 걸쳐 혁신적인 변화를 이끌고 있다. 멀티모달 AI는 생성형 AI와 함께 최근 AI 혁신의 선봉에 서 있으며, 현재 구글(Google) 제미니(Gemini), 오픈AI(OpenAI) Chat-GPT 모두 멀티모달 기능을 제공한다.

멀티모달 AI와 생성형 AI(Generative AI)는 서로 보완적인 역할을 한다. 멀티모달 AI는 다양한 데이터 소스를 통합하고 분석하며, 생성형 AI는 이를 바탕으로 새로운 콘텐츠를 다양한 양식으로 생성한다. 이 두기술은 최첨단 AI의 서로 다른 측면이라고 이해할 수 있다. 이번 원고에선 멀티모달 AI의 기본 개념, 기술 동향, 응용 분야, 미래 전망에 대해 살펴보고자 한다.

2. 멀티모달 AI의 기본 개념

2.1 멀티모달 데이터

데이터 모달이란 정보를 표현하고 전달하는 방식 또는 매체를 의미한다. 즉, 멀티모달 데이터는 서로 다른 유형의 정보를 포함하는 데이터다. 텍스트, 이미지, 오디오, 비디오 등 다양한 형태가 있으며, 각 데이터 유형은 고유한 특성과 정보를 담고 있다.

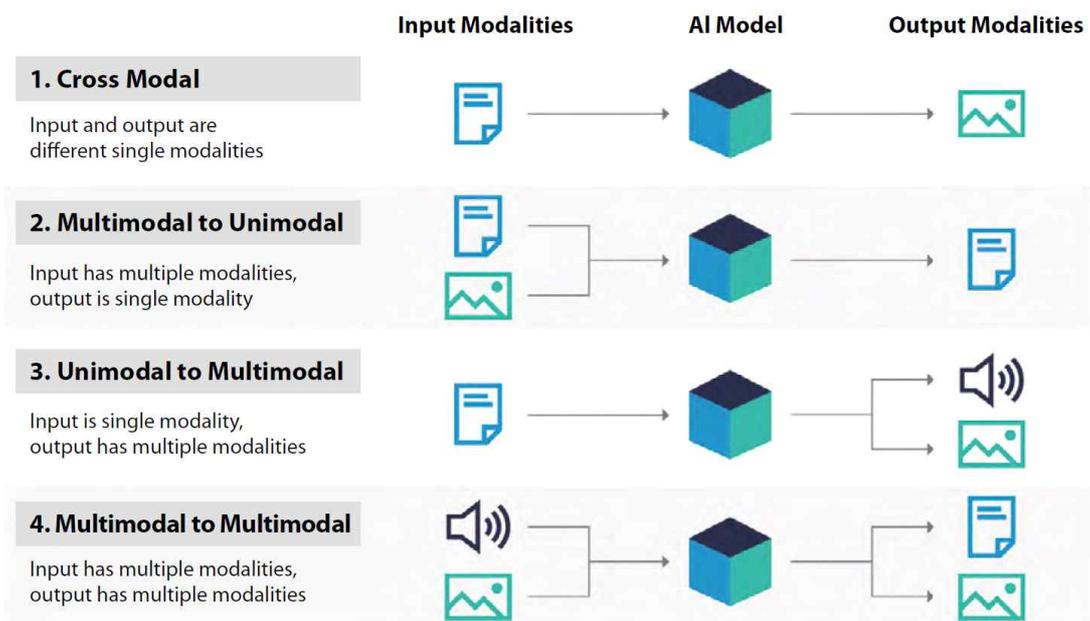
- 텍스트: 논문, 소셜 미디어 게시물, 뉴스 기사 등 자연어로 작성 정보를 포함. 감정 인식, 내용 작성 및 요약, 주제 분류 등 다양한 언어 처리 작업에 활용.
- 이미지: 사진, 그림, 차트 등 시각적 정보를 포함. 객체 인식, 장면 이해, 이미지 분류 등의 컴퓨터 비전 작업에 활용.
- 오디오: 음성 녹음, 음악, 소리 신호 등 청각적 정보를 포함. 음성 인식, 감정 분석, 소리 분류 등의 작업에 사용.
- 비디오: 시간에 따라 변화하는 시각 및 청각 정보를 포함한 동영상 데이터를 의미. 동작 인식, 장면 분석, 행동 예측 등 복잡한 작업에 활용.
- 코드: 주어진 입력과 맥락에 따라 특정한 작업을 수행하기 위한 명령어들의 집합. 파이썬 (Python), 자바(Java) 등 다양한 프로그램 언어를 지원.

이 외에도 온도, 습도, 압력 등의 센서 데이터도 멀티모달 데이터에 포함될 수 있다.

2.2 멀티모달 AI

멀티모달 AI는 멀티모달 데이터를 동시에 처리하고 분석함으로써 더 풍부하고 정확한 정보를 제공한다. 예를 들어, 제품에 대한 리뷰를 분석할 때, 텍스트 리뷰와 함께 제품 이미지, 사용 동영상, 추천 버튼 등을 함께 활용하면 더 종합적인 분석 결과를 얻을 수 있다. 또한, 미드저니 (Midjourney), 달리(DALL-E), 소라(Sora)의 경우 사용자가 텍스트로 상황을 기술하면, 이를 기반으로 이미지나 동영상을 생성한다. [그림 1]은 멀티모달 AI의 분류다[2].

Taxonomy With Different Types of Multimodal AI



출처: Ramos, L., et al., Innovation Insight: Multimodal AI Explained, Gartner, G00798532. 2023, 1-13.

[그림 1] 멀티모달 AI의 분류

최근엔 입력과 출력 양식이 다른 크로스 모달 형태가 주목받고 있다. 그 외에도, 단일 양식 입력에서 다중 양식 출력을 생성하거나, 혹은 그 역이거나, 아니면 다중 양식 입력을 기반으로 다중 양식 출력을 제공하는 형태도 있다.

2.3 멀티모달 데이터의 통합

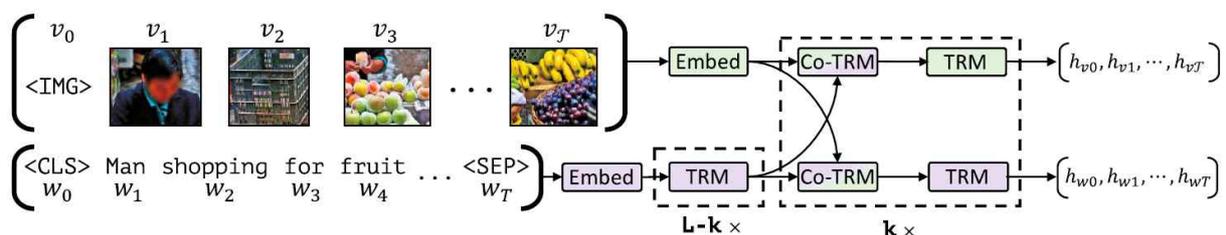
멀티모달 AI의 핵심은 멀티모달 데이터 통합 기술이다. 멀티모달 데이터 통합은 다양한 데이터 유형을 결합해 종합적으로 분석하고 이해할 수 있도록 한다. 이를 통해 단일 데이터 유형만으론 분석이 어려운 복잡한 정보를 다양한 각도에서 분석할 수 있다.

여러 데이터 유형을 결합하면 더 정확한 예측과 결정을 내릴 수 있다. 예를 들어, 미묘한 감정을 인식하는 데 있어, 텍스트와 이미지를 결합해 감정 분석을 수행하면 더 신뢰성 있는 결과를 얻을 수 있다. 또한, 다양한 데이터 유형을 통합하면 상황에 대한 더 깊고 포괄적인 이해가 가능해진다. 이는 특히 의료, 자율주행 등 복잡한 시스템에서 중요한 특징이다.

또한, 이미지 데이터가 부족한 경우, 확보된 텍스트나 음성 데이터를 활용해 부족한 데이터를 보완할 수 있으므로, 데이터 부족 문제에도 강한 특징을 갖는다. 데이터 통합은 일반적으로 아래와 같은 절차를 따른다.

- 데이터 전처리: 다양한 데이터 유형을 통합하기 위해선 먼저 각 데이터 유형에 적합한 전처리 과정을 거친다. 텍스트의 경우 토큰화, 이미지나 비디오는 해상도 조정, 프레임 추출 등의 과정이 필요하다.
- 특징 추출: 각 데이터 유형에서 중요한 특징을 추출한다. 텍스트 데이터에선 단어 임베딩 값을, 이미지 데이터에선 CNN을 통해 추출된 특징 맵을 사용할 수 있다.
- 공통 표현 학습: 다양한 데이터 유형을 공통된 표현 공간으로 매핑해 결합한다. 이를 위해 멀티모달 딥러닝 모델을 사용한다. 이러한 모델은 서로 다른 양식 간의 관계를 학습하고, 공통된 표현을 통해 데이터를 결합한다.
- 통합 분석: 공통 표현을 통해 결합된 데이터를 기반으로 예측·분류 등의 분석을 수행한다.

ViLBERT(Vision-and-Language BERT)[3]는 FAIR(Facebook AI Research)에서 개발한 멀티모달 AI 모델로, 텍스트와 이미지를 함께 처리해 이들의 상호관계를 이해한다. ViLBERT는 BERT(Bidirectional Encoder Representations from Transformers) 구조를 기반으로, 이미지와 텍스트 데이터를 동시에 처리하고 상호작용을 학습할 수 있도록 확장한 모델이다. [그림 2]는 ViLBERT가 데이터 통합을 하는 과정으로, 앞서 언급한 데이터 통합 절차를 아주 잘 보여준다.



출처: Lu, J., et al. (2019). ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and-Language Tasks. 33rd Conference on Neural Information Processing Systems (NeurIPS)[1]

[그림 2] ViLBERT 학습 방법

이를 보면, 크게 두 단계 학습 과정을 거치는 것을 알 수 있다. 첫 번째 과정은 텍스트, 이미지 등 개별 양식에 대한 학습, 두 번째는 멀티모달 데이터를 이용해 두 양식을 하나의 표현 공간으로 임베딩하는 학습이다. 이 모델은 텍스트와 이미지 간 상호관계를 학습하기 위해 두 개의 분리된 트랜스포머(Transformer) 모듈을 사용하고, 양식 간 정보 교환을 위해 코-어텐션(Co-attention) 메커니즘을 적용한다. 이를 통해 ViLBERT는 텍스트와 이미지가 통합된 표현을 생성하며, 다양한 멀티모달 작업에서 효과적으로 활용된다. 즉, 데이터 양식별로 전처리하고 특징을 추출해 임베딩을 생성한다. 그리고 이렇게 생성된 양식별 임베딩과 멀티모달 데이터를 이용해 이미지와 텍스트 양식에 대한 공통의 표현을 학습함으로써 데이터를 통합한다.

이렇게 통합된 데이터 표현을 기반으로, 이미지 캡셔닝, 이미지-텍스트 매칭과 같은 다운스트림 작업이 수행된다. AI에서 '임베딩(Embedding)'은 고차원 데이터를 연속적인 저차원 벡터 공간으로 변환하는 기술을 의미한다. 이러한 변환의 목적은 데이터의 의미나 유사성을 보존하면서, 모델이 처리하기 쉽게 만드는 것이다. 데이터 통합의 실제적 방법과 절차는 모델별로 다양하며, 위에서 언급한 방법은 대표적인 사례다.

3. 멀티모달 AI 기술 동향 및 주요 응용 분야

3.1 최신 기술 및 발전

멀티모달 AI는 다양한 데이터 유형을 통합해 더 풍부한 이해와 예측을 가능케 하는 기술로, 최근 몇 년동안 빠른 속도로 발전하고 있다. 이러한 발전은 새로운 모델 구조, 효율적인 학습 방법, 대규모 데이터셋의 활용 등을 포함한다.

3.1.1 모델 구조

트랜스포머 모델은 자연어 처리에서 혁신을 가져왔으며, 현재 이를 멀티모달 데이터에 적용한 모델들이 개발되고 있다. ViLBERT, 클립(CLIP)[4], 달리[5]는 모두 트랜스포머를 기반으로 한다. ViLBERT와 오픈AI의 클립은 이미지와 텍스트 쌍을 사용해 훈련된 모델이며, 특히 이미지와 텍스트의 의미 이해에 탁월하다. 이는 텍스트와 이미지의 임베딩을 공통된 공간으로 맵핑해 다양한 다운스트림 작업에서 활용될 수 있다. 오픈AI의 달리는 텍스트 설명을 기반으로 이미지를 생성할 수 있는 모델로, 텍스트와 이미지 간 복잡한 관계를 학습해 창의적인 콘텐츠 생성이 가능하다. 텍스트 이해를 위해선 트랜스포머 기반 모델을, 이미지 생성을 위해선 VAE(Variational Autoencoder) 기반 모델을 사용했다. 그 외 GAN을 비롯한 다양한 딥러닝 기술에 기반한 모델들도 출시됐다. 이미지바인드(ImageBind)[6]는 FAIR에서 개발한 멀티모달 AI 모델로, 6개 양식(텍스트, 이미지/비디오, 오디오, 깊이, 열, IMU) 데이터에 대한 공통된 표현을 학습한다.

3.1.2 효율적인 학습 방법

멀티모달 AI의 효율적인 학습 방법은 계산량을 줄이거나, 부족한 양식의 데이터를 보강하거나, 계산된 모델을 효율적으로 재활용함으로써 멀티 모달 데이터를 효과적으로 통합하는 것과 관계된다.

- 사전 학습된 양식별 모델 사용: 사전 학습된 양식별 모델을 바탕으로 초기 표현을 생성해 활용한다. 텍스트의 경우 트랜스포머 또는 BERT, 이미지의 경우 ResNet 또는 ViT을 사용해, 각 양식에 대한 강력한 기초 표현을 얻을 수 있다[4]. 물론 양식별 인코딩을 위해서 다른 모델도 활용가능하다.
- 공통 임베딩 공간 학습: 양식별 데이터를 공통된 벡터 공간으로 변환해, 이들 간의 상호작용을 학습한다. 서로 다른 유형의 데이터를 비교하고 결합하기 쉬워진다[3][4][6].
- 데이터 증강 및 정규화: 각 양식에 대해 다양한 데이터 증강 기법과 정규화를 적용해 모델이 다양한 상황에서 잘 작동하도록 학습한다.
- 지식 증류(Knowledge Distillation): 큰 모델(Teacher Model)에서 작은 모델(Student)로 지식을 전이해 모델의 효율성을 높이는 방법을 사용함으로써 연산 비용을 줄이고 성능을 유지하는 데 유리하다[7].
- 전이 학습(Transfer Learning): 다양한 양식에서 사전 학습된 모델을 활용해 새로운 작업에 적응시킨다. 이는 특히 데이터가 부족한 상황에서 유용하게 사용될 수 있다[8].
- 공동 학습(Joint Learning): 다양한 양식의 데이터를 동시에 학습해, 각 양식의 정보를 최대한 활용할 수 있다. 이를 통해 모델은 더 풍부한 표현을 학습할 수 있으며, 다양한 입력 유형을 효과적으로 처리할 수 있다[9].

3.1.3 대규모 데이터셋의 활용

멀티모달 AI 기술 개발을 위해, 다양한 대규모 데이터셋이 활용되고 있다. 대표적으로 마이크로소프트(Microsoft) COCO(Common Objects in Context), LSMDC(Large Scale Movie Description Challenge) 등이 있다.

3.2 대표적인 솔루션과 응용 분야

멀티모달 AI의 발전을 이끄는 대표적인 솔루션과 응용 분야는 <표 1>과 같다.

<표 1> 멀티모달 AI 기술 개발에 활용되는 대규모 데이터셋

이름	모달리티	설명	용도
MS COCO	이미지, 텍스트(객체 클래스, 캡션), 좌표 등	다양한 일상적인 장면에서 객체를 인식하고 캡션을 생성하는 데 사용되는 데이터셋	이미지 캡셔닝, 객체 검출, 장면 이해 등
Visual Genome	이미지, 텍스트(QA, 설명)	이미지의 시각적 이해를 돕기 위해 객체, 속성, 관계 정보를 포함	시각적 질의응답, 이미지 캡셔닝, 멀티모달 검색 등
LSMDC	비디오, 텍스트(캡션)	영화 클립과 해당 클립에 대한 자연어 캡션을 포함하는 데이터셋	비디오 캡셔닝, 비디오-텍스트 매칭, 장면 이해 등
HowTo100M	비디오, 텍스트(캡션)	유튜브 영상과 해당 영상의 자막을 활용해 다양한 작업을 배우는 데 사용되는 대규모 멀티모달 데이터	행동 인식, 동작 예측, 자막 생성, 멀티모달 학습 등

3.2.1 대표적인 솔루션

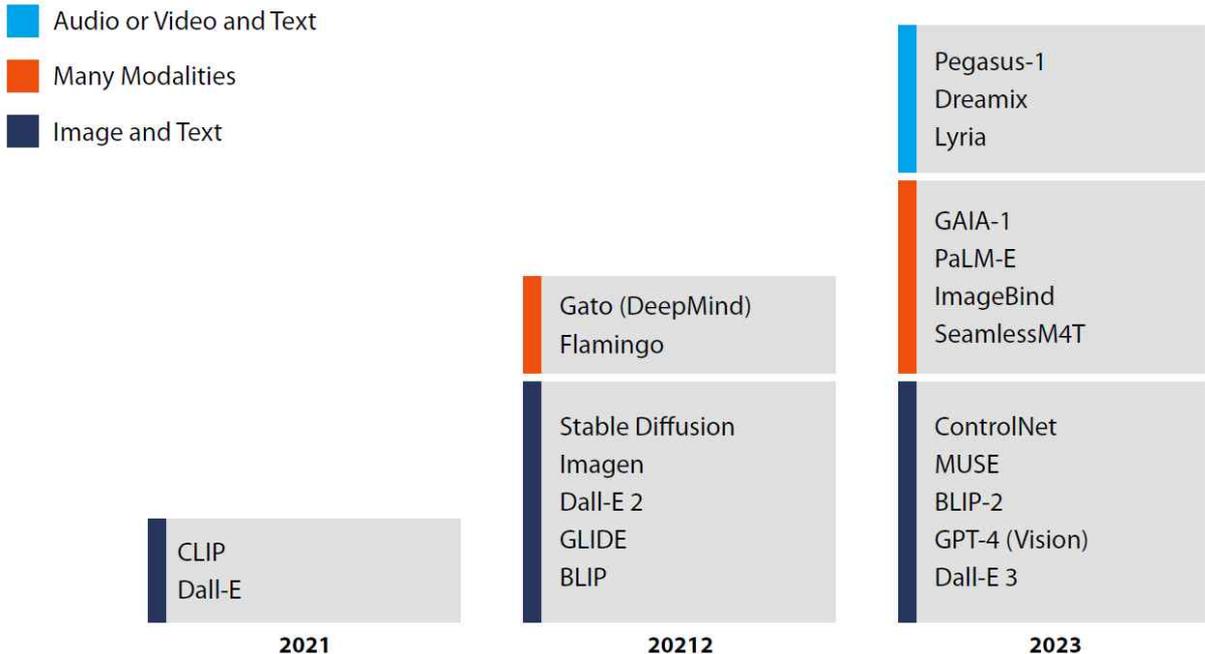
대표적인 멀티모달 솔루션은 <표 2>와 같다.

<표 2> 멀티모달 솔루션

이름	모달리티	설명	용도
구글 제미니	텍스트, 이미지, 오디오, 비디오, 코드	다양한 모달리티 데이터를 통합해 처리하고 분석할 수 있는 멀티모달 AI 솔루션	검색 엔진, 추천 시스템, 가상 비서, 콘텐츠 생성 등
오픈AI Chat-GPT	텍스트, 이미지, 코드	자연어 처리와 텍스트 기반 대화를 지원하는 AI 모델로, 이미지 이해 및 설명 생성 기능을 포함해 멀티모달 데이터 처리 가능	고객 지원, 콘텐츠 생성, 교육 도구, 대화형 AI 등
마이크로소프트 Azure Cognitive Services	텍스트, 이미지, 오디오, 비디오	다양한 모달리티의 데이터를 처리해 인식, 번역, 분석 기능을 제공하는 AI 솔루션	멀티미디어 콘텐츠 분석, 텍스트 분석, 음성 인식 등
아마존 AI Services	텍스트, 이미지, 음성, 비디오	다양한 모달리티 데이터를 처리하고 분석하는 멀티모달 AI 기능을 제공하는 클라우드 기반 플랫폼	멀티모달 표현 학습, 번역, 멀티모달 인지 등

이 외에도 [그림 3]과 같은 멀티모달 기술들이 출시돼 있다. 2021년 클립, 달리를 시작으로 2023년까지 다양한 멀티모달 AI 기술들이 출현했다.

Multimodal AI Models Growth and Evolution



출처: Ramos, L., et al., Innovation Insight: Multimodal AI Explained, Gartner, G00798532. 2023, 1-13

[그림 3] 멀티모달 AI 모델의 성장

3.2.2 응용 분야

멀티모달 AI는 다양한 산업에서 혁신적인 응용 사례를 제공하고 있다.

• 엔터테인먼트

트웰브 랩스(Twelve Labs)는 멀티모달 AI를 활용해 엔터테인먼트 분야에서 콘텐츠를 생성하고 있다[10]. 이 기술은 텍스트, 이미지, 오디오, 비디오 데이터를 통합해 비디오를 분석하고 생성한다. 예를 들어, 트웰브 랩스의 페가수스(Pegasus), 마렌고(Marengo) 모델은 비디오 내용을 요약하고, 특정 장면을 검색하며, 설명 텍스트를 생성하는 기능을 제공한다. 이러한 기술은 포스트 프로덕션 워크플로우 최적화, 사용자 맞춤형 콘텐츠 추천, 비디오 분류·검색 등 다양한 분야에서 활용된다. 이러한 멀티모달 AI 기술은 영화, TV, 디지털 콘텐츠 제작에서 더 나은 품질의 콘텐츠를 더 빠르고 효율적으로 생산할 수 있게하며, 엔터테인먼트 산업 전반에 걸쳐 큰 영향을 미치고 있다.

• 스마트 시티

전력 사용 데이터, 기상 데이터, 건물 에너지 소비 패턴 등을 통합해 에너지 효율을 최적화할 수 있다. 예를 들어, 실시간 기상 데이터와 건물 에너지 소비 데이터를 결합해 냉난방 시스템을 자동으로 조절함으로써 에너지를 절약할 수 있다. 또한, CCTV 영상, 소셜 미디어 데이터, 경찰 보고서 등을 통합해 범죄 예측 및 대응을 강화할 수도 있다. 가령 특정 지역에서의 범죄 발생 패턴을 분석해 경찰의 순찰 경로를 최적화하는 것이 가능하다.

• 금융

멀티모달 AI는 거래 기록, 고객 프로필, 소셜 미디어 데이터 등을 결합해 금융 사기를 더 정확하게 탐지하고 예방한다. 예를 들어, 비정상적인 거래 패턴과 소셜 미디어에서의 의심스러운 구와 불만을 신속하게 파악하고 대응할 수 있다. RoBERTa-Fin 모델[11]은 뉴스 헤드라인, 재무 보고서와 같은 텍스트 데이터와 주식 데이터, 경제 데이터 등의 수치 데이터를 융합해 금융시장 예측 및 분석의 정확도를 높인 혁신적인 모델이다.

4. 기술적 과제와 미래 전망

멀티모달 AI는 다양한 데이터 유형을 통합해 고도의 분석과 예측을 가능하게 하지만, 이 과정에서 여러 기술적 과제에 직면한다. 다양한 데이터 유형을 수집하고 전처리하는 과정은 매우 복잡하고 시간을 많이 소모한다. 더불어 데이터 통합과 동기화 또한 큰 도전과제다. 이러한 모델은 해석 가능성과 공정성 문제로 인해 신뢰성과 투명성 측면에서 어려움을 겪고 있으며, 높은 계산 비용은 지구 환경의 지속가능성 차원에서 위협 요인이 되고 있다.

그럼에도 불구하고, 멀티모달 AI는 다양한 분야에서 혁신적 변화를 불러올 것으로 기대된다. 특히 모델 통합 및 상호 운용성, 실시간 처리 능력 향상, 에너지 효율성, 지속가능성 등이 중요한 발전 방향이 될 것으로 보인다. 향후 자율주행, 로봇공학, 스마트 시티 분야 등에서 다양한 모달

리티 데이터를 효과적으로 결합하고 실시간으로 처리하는 기술이 발전할 것이다. 이를 바탕으로 개인화된 맞춤형 서비스 제공이 가능해지며, 의료, 교육, 엔터테인먼트 분야에서 큰 변화를 불러올 전망이다. 또한, 윤리적 문제와 규제 환경에 관한 연구는 안전하고 공정한 기술 사용을 보장하는 데 기여할 것이다.

5. 맺음말

멀티모달 AI의 핵심은 다양한 양식의 데이터를 공통의 표현 공간으로 투사시키는 기술이다. 텍스트, 이미지, 비디오 등 서로 다른 양식의 데이터를 같은 공간에 표현함으로써 형식 이전에 데이터가 갖는 의미, 관계를 좀 더 정확하고 폭넓게 이해할 수 있다.

이러한 기술적 시도는 향후 AI 기술의 응용 범위와 인식 범위를 매우 넓힐 수 있는 잠재력을 갖고 있다. 서로 다른 양식에 기반한 검색(Cross-Modal Retrieval)은 물론, 음성·이미지 데이터에 대한 연산만으로 새로운 음성·이미지 데이터를 생성할 수 있게 된다.

다양한 미디어 생성이 다양한 방식으로 가능해짐에 따라, 창의력을 기반으로 멀티 모달 예술 작품을 용이하게 창작할 수 있는 시기가 도래할 것이다. 세잔의 시도가 있는 그대로의 그림을 그리던 시기에서 이를 해체하는 입체파라는 새로운 장르를 열게 한 것과 마찬가지로, 멀티모달 AI 기술은 AI 서비스의 새로운 장르를 열 모멘텀이 될 것이다.

※ 본 연구는 산업통상자원부(MOTIE)와 한국에너지기술 평가원(KETEP)의 지원을 받아 수행한 연구 과제이다. (No.20202020900290)

[참고문헌]

- [1] 김말희 외, 인간을 닮은 AI, 멀티모달 AI 기술 동향, 2136호, 2-16.
- [2] Ramos, L., et al., Innovation Insight: Multimodal AI Explained, Gartner, G00798532. 2023, 1-13.
- [3] Lu, J., et al., ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and-Language Tasks. 33rd Conference on Neural Information Processing Systems(NeurIPS), 2019, 1-11.
- [4] Radford, A., et al., Learning Transferable Visual Models From Natural Language Supervision, Proceedings of the 38th International Conference on Machine Learning, PMLR 139, 2021, 1-16.
- [5] Ramesh, A., et al., Zero-Shot Text-to-Image Generation, arXiv:2102.12092v2, 2021, 1-20.
- [6] Girdhar, R., et al., IMAGEBIND: One Embedding Space To Bind Them All, The IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, 15180-15190.
- [7] Hinton, G. et al., Distilling the Knowledge in a Neural Network, arXiv:1503.02531v1, 2015. 1-9.
- [8] Wang, Y., et al., VideoAdviser: Video Knowledge Distillation for Multimodal Transfer Learning, arXiv:2309.15494v1, 2023, 1-17.
- [9] Fritsch S.G., et al., MuJo: Multimodal Joint Feature Space Learning for Human Activity Recognition, arXiv:2406.03857v1, 2024, 1-13.

[10] Twelve Labs, <https://www.twelvelabs.io/>

[11] Das S., et al., Context, Language Modeling, and Multimodal Data in Finance, The Journal of Financial Data Science, 2021, 1-15.

※ 출처: TTA 저널 제214호