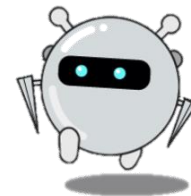


CV&NLP 연구동향



진행자 : 멘토 현시은

CV&NLP 연구동향

- 컴퓨터비전(CV)
- 자연어처리(NLP)
- 그 외(etc)

컴퓨터비전

- 컴퓨터비전

자연어처리

그 외

- 생성형 AI의 시대

- 이미지 분야에서 가장 뜨거운 연구 분야는 단연 이미지와 비디오 합성 분야
- 이미지와 비디오 합성 분야는 학계에서의 연구뿐만 아니라 상업적으로도 응용 가능성이 매우 큰 분야이다.

- 컴퓨터비전

자연어처리

그 외

- GAN으로 시작해서 Diffusion에 이르기까지
 - 딥러닝을 이용하여 새로운 이미지를 '창조'할 수 있다는 것은 엄청난 파급력을 가져왔다.
 - GAN은 아주 유용한 알고리즘이지만, 최근 이미지 생성 연구에서는 Diffusion을 활용한 모델이 많이 연구되고 있다.
 - Diffusion 이론을 이용하여 퀄리티가 좋은 이미지 생성 모델들이 경쟁적으로 출시중!

- 컴퓨터비전

자연어처리

그 외

- 이제는 2D 이미지에서 안주하지 않는다.

OpenAI에서 SORA를 출시하여 딥러닝을 이용하여 비디오까지도 생성할 수 있음을 보였다.



<https://openai.com/index/video-generation-models-as-world-simulators/>

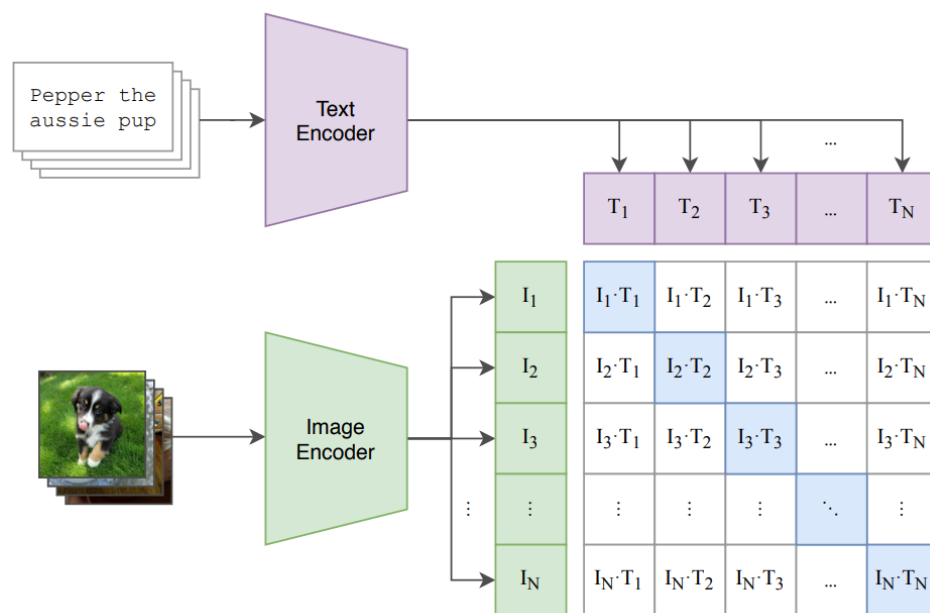
• 컴퓨터비전

자연어처리

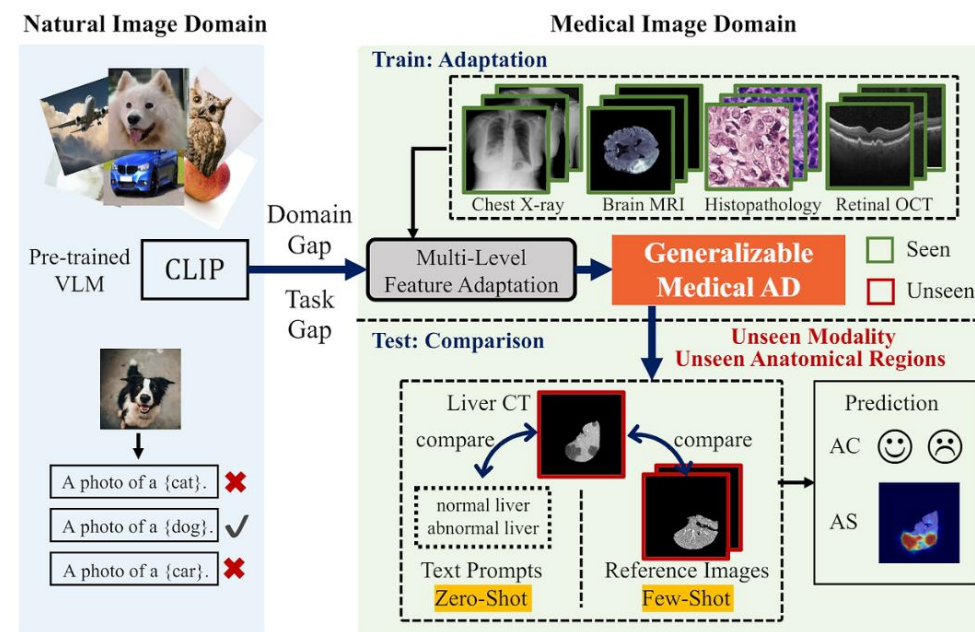
그 외

• Visual-Language Model (VLM)

- 여러가지 Modality를 하나의 모델로 다루는 연구가 지속됨에 따라서, 최근 Visual-Language Model (VLM)이 좋은 성능을 보이고 있다.



<https://arxiv.org/pdf/2103.00020>



<https://arxiv.org/pdf/2403.12570>

- 컴퓨터비전
자연어처리
그 외

- 3D image, Segmentation, Object Detection 등의 분야도 활발히 연구



<https://ai.meta.com/blog/segment-anything-2/>

"Elon musk, using a laptop"



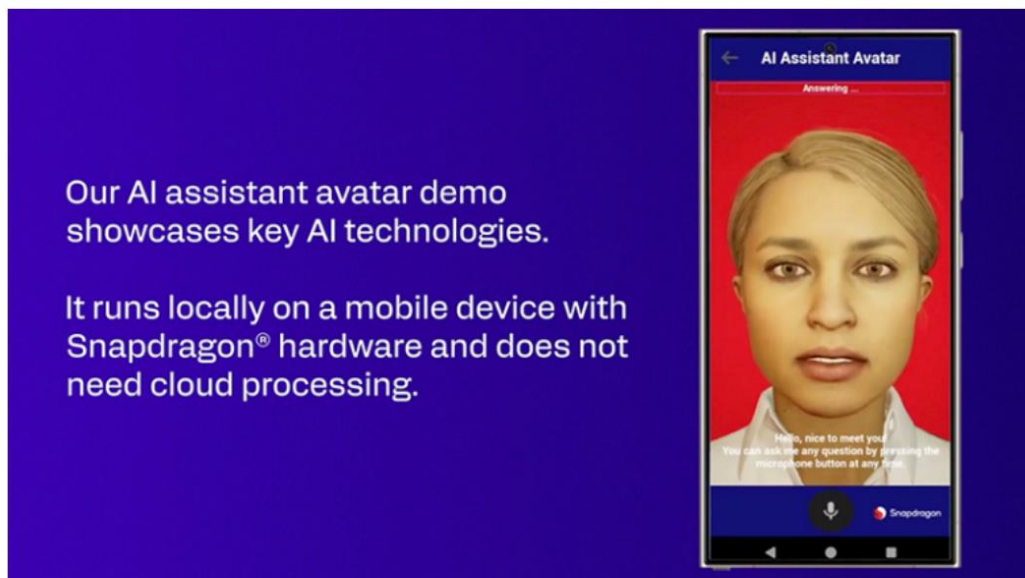
<https://github.com/tyhuang0428/DreamControl>

- 컴퓨터비전

자연어처리

그 외

- 이제는 성능만 좋아서는 안된다!



온디바이스 AI 어시스턴트는 오디오 기반 3D 아바타를 사용합니다.

- 퀄컴은 LoRA 어댑터를 활용해 안드로이드 기반 스마트폰에서 실행되는 스테이블 디퓨전(Stable Diffusion)을 시연했다.

[출처] <https://blog.naver.com/qualcommkr/223510097268>

- AI의 시장이 엄청나게 커졌고, 이제는 연구 레벨을 넘어 상업적으로도 개발이 되고 있는 만큼 '실제로 사용할 수 있는지' 여부도 중요해짐
- 성능은 좋지만, 연구 시설(즉, 엄청난 컴퓨팅 파워가 지원되는 상황)에서만 사용 가능하다면 이는 실용적이지 못함
- 다양한 최적화 기술을 사용하여 온디바이스에서도 사용할 수 있는 이미지 생성 AI 또한 활발하게 연구중!

자연어처리

컴퓨터비전

- 자연어처리

그 외

- Transformer의 등장

- 2017년에 등장한 Transformer를 기반으로 하는 엄청나게 다양한 Language Model이 등장함
- 그 중 가장 성공적이라고 할 수 있는 GPT 시리즈의 굉장한 성과를 시작으로, 수많은 LLM이 탄생하였다. (Llama, Gemini, phi, Claude ……)

컴퓨터비전

- 자연어처리

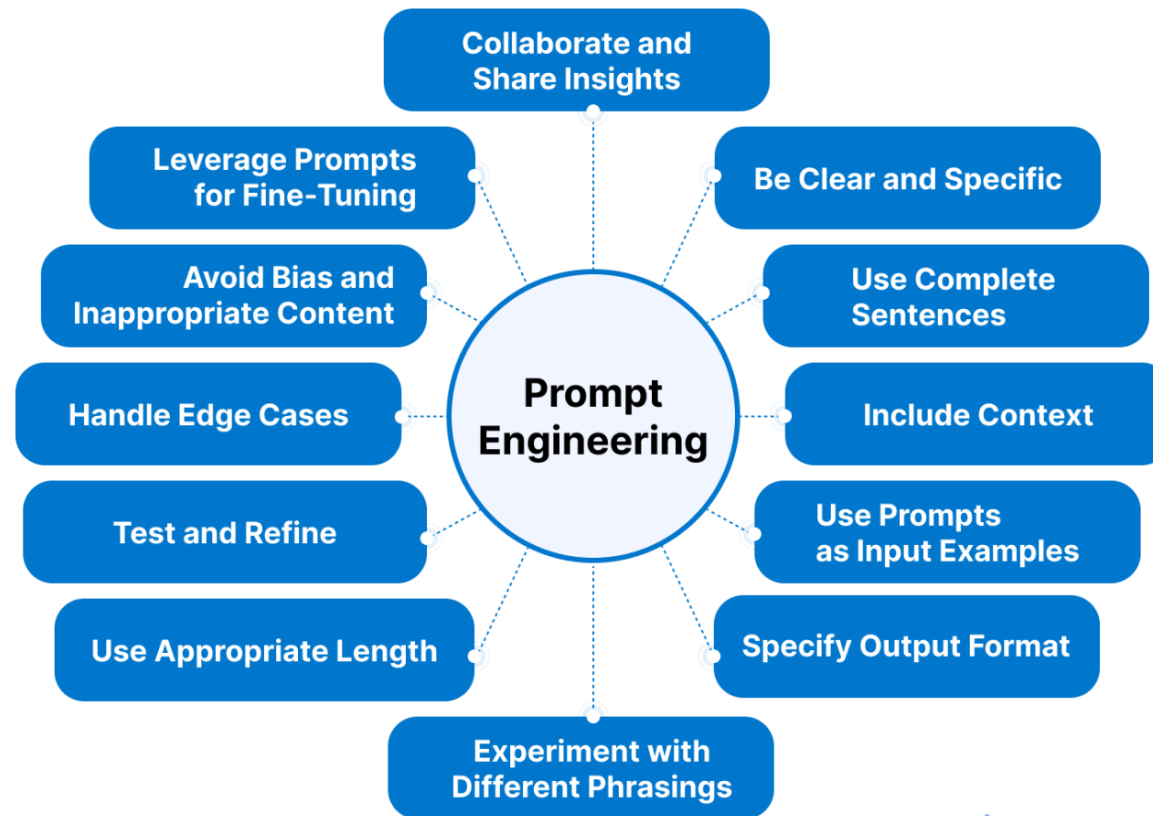
그 외

- LLM의 시대

- 엄청난 LLM의 퍼포먼스는 전세계를 열광시켰다.
- 전세계의 글로벌 기업들이 수 조원의 돈을 투자하면서 LLM 개발에 앞장서고 있다.
 - Google, Meta, Microsoft, OpenAI, NVIDIA 등...

과연 일반 연구 단체가 이들의 자본력을 뛰어넘는 연구 성과를 낼 수 있을까...?

- LLM의 능력은 아직도 전부 밝혀지지 않았다...!
 - 프롬프트 엔지니어링의 등장.
 - 단순히 프롬프트를 잘 입력하는 것만으로도 성능이 달라지는 현상을 '발견'함



컴퓨터비전

- 자연어처리

그 외

- 이제는 하드웨어가 부족하다...!
 - LLM은 일반 컴퓨터로는 사용하는 것이 불가능하다.
 - LLM 서비스를 위해서는 수천 억원 규모의 컴퓨터 자원이 필요하다.
(NVIDIA가 엄청난 성장을 하게 된 원인)

이게 정말 바람직한 방향일까?

- sLLM의 등장
 - LLM의 성능을 유지하면서 하드웨어에 최적화시키는 연구가 활발히 진행중
 - 즉, LLM을 api로 원격 서버에서 사용하는 것이 아니라, 나의 로컬 컴퓨터에서 사용할 수 있게 하는 온디바이스 LLM을 개발하려고 한다.
 - 예 : Gemini-nano, phi-3 mini,
 - 이를 위해서 Quantization, PIM, NPU, DPU 등 다양한 AI friendly 하드웨어 연구도 활발히 진행중이다.

컴퓨터비전

• 자연어처리

그 외

- sLLM의 등장

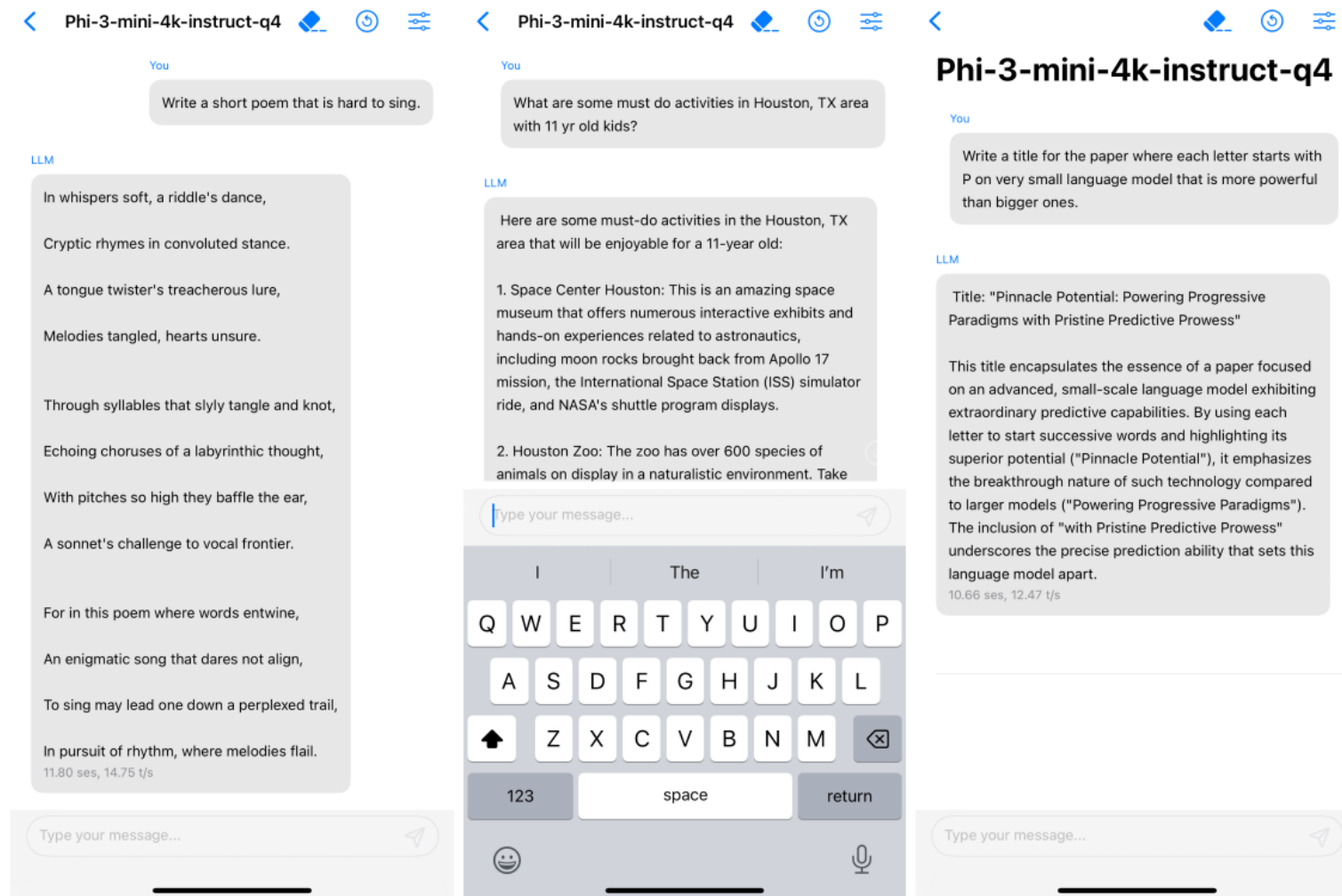
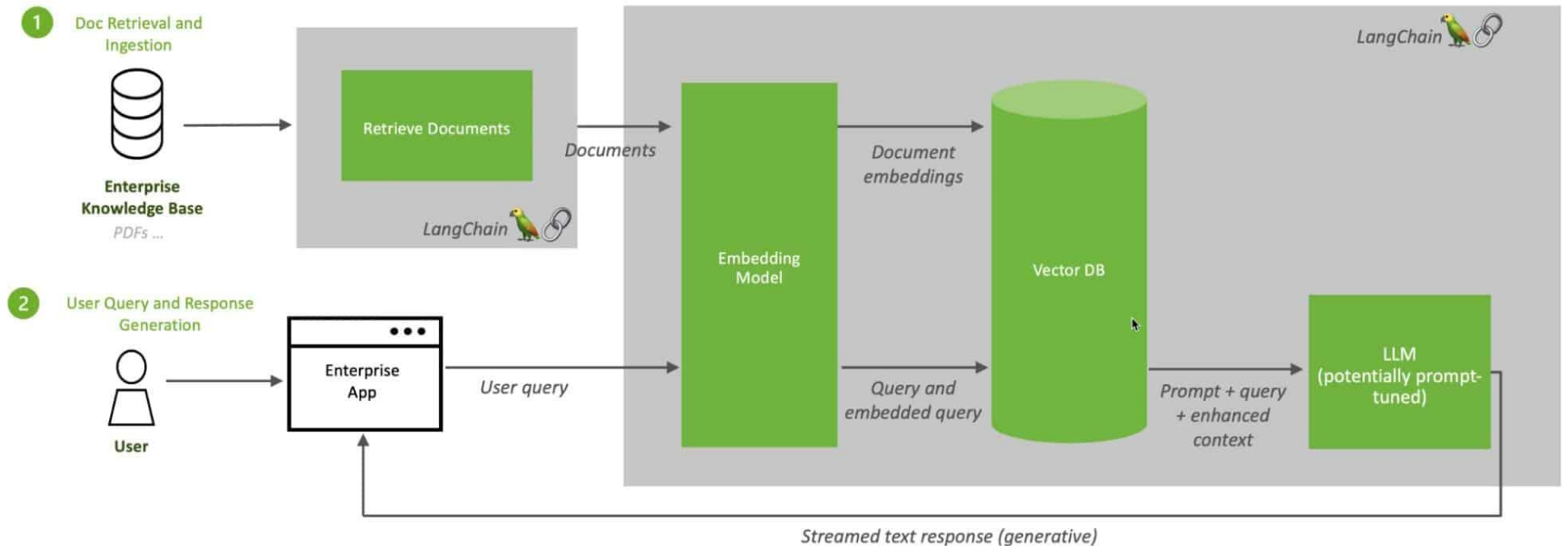


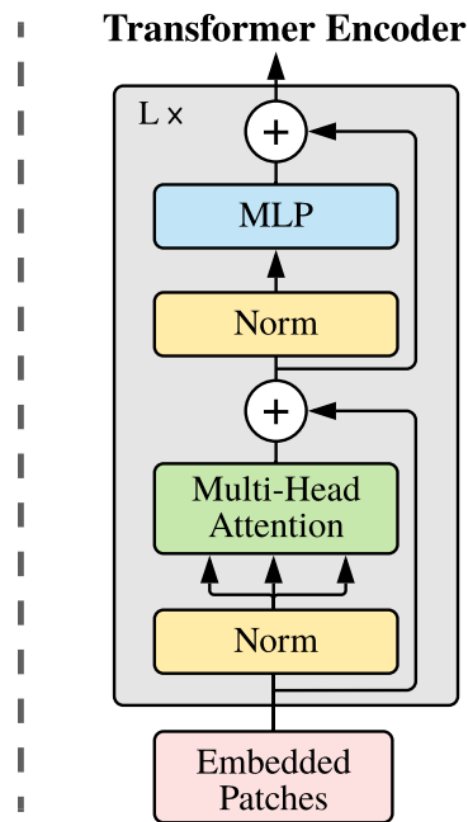
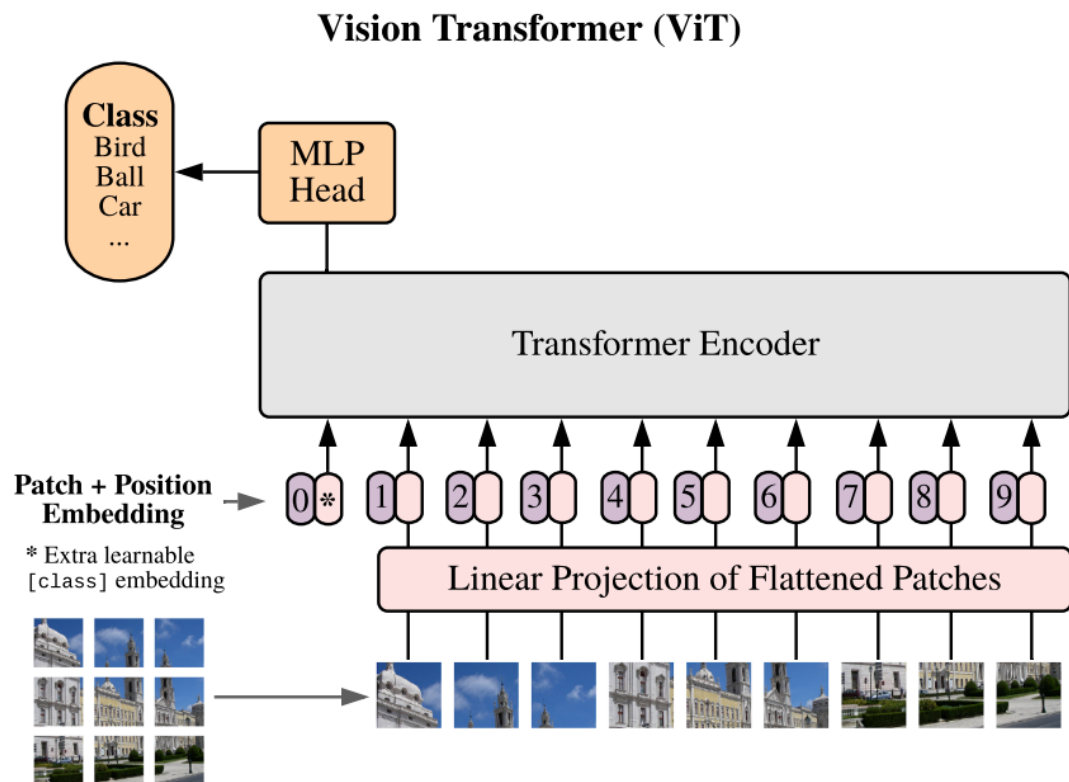
Figure 2: 4-bit quantized **phi-3-mini** running natively on an iPhone with A16 Bionic chip, generating over 12 tokens per second.

- LLM을 위해서는 하드웨어만 필요한 것이 아니다! - RAG의 등장
- LLM을 DataBase(DB)와 함께 사용하여 더 성능을 높이는 연구의 등장

Retrieval Augmented Generation (RAG) Sequence Diagram



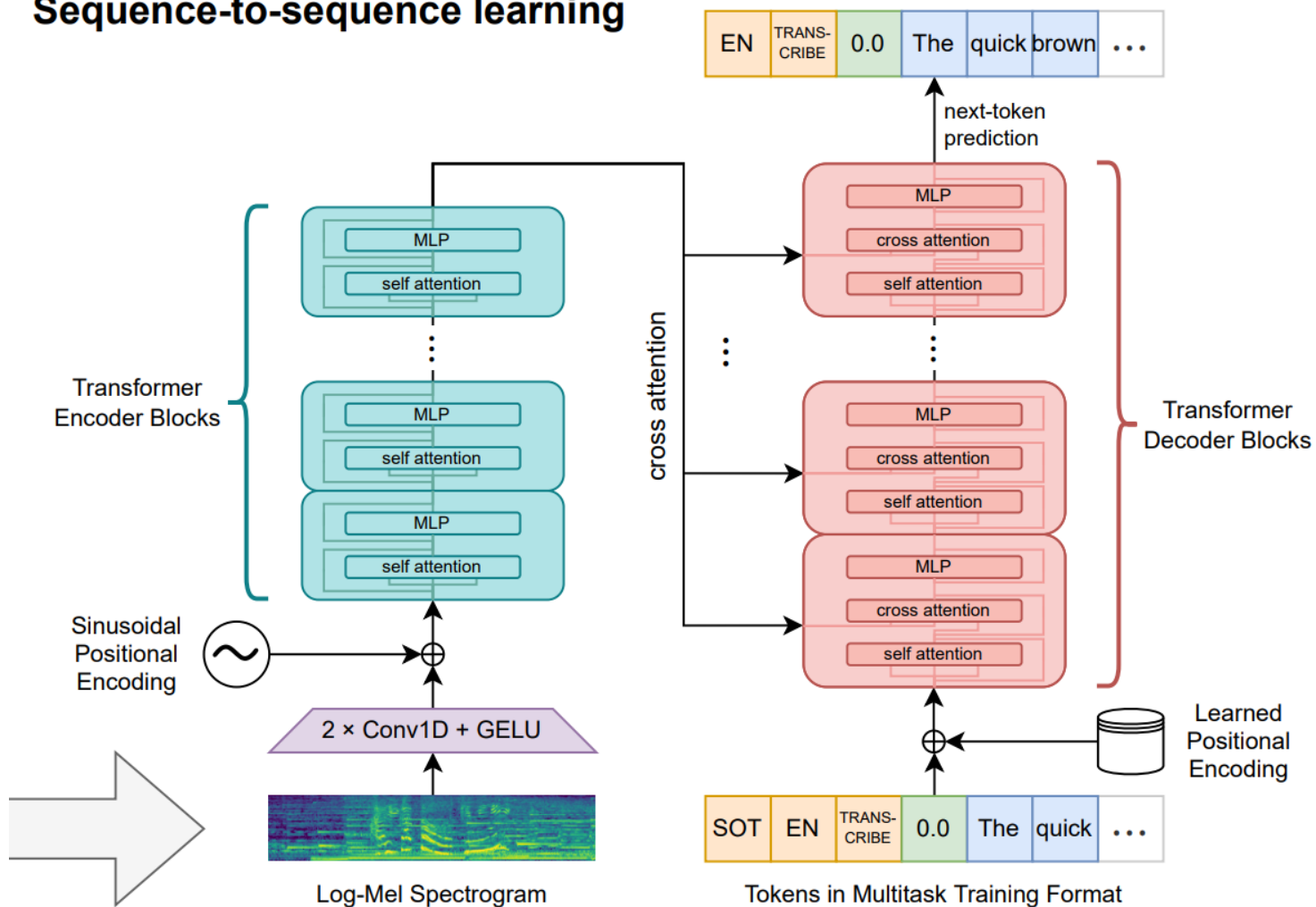
- Language Model의 방법론은 다른 분야에도 다양하게 쓰인다(in CV)



<https://arxiv.org/abs/2010.11929>

- Language Model의 방법론은 다른 분야에도 다양하게 쓰인다(in ASR)

Sequence-to-sequence learning



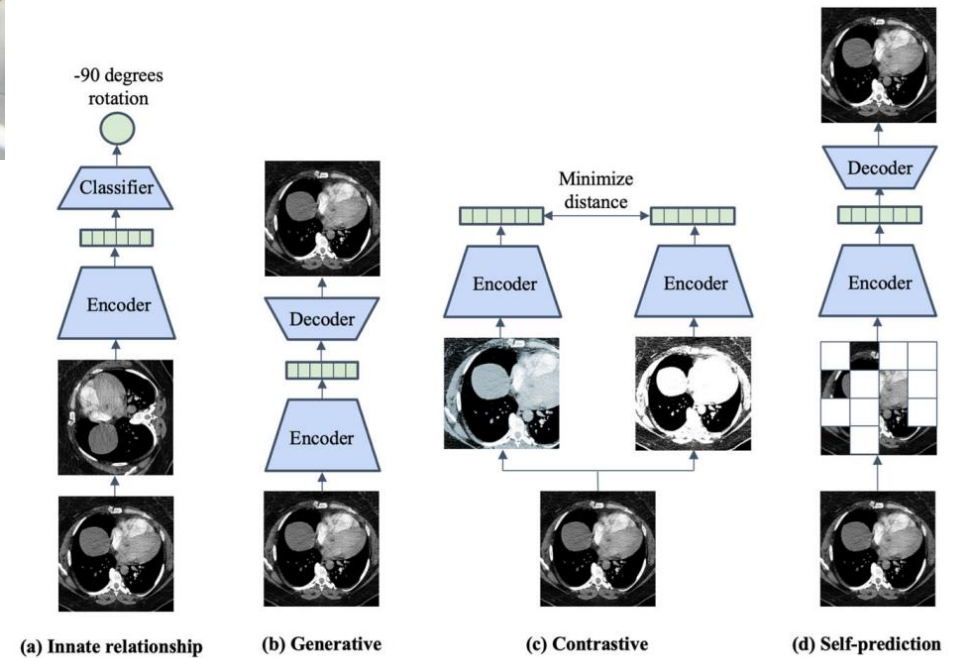
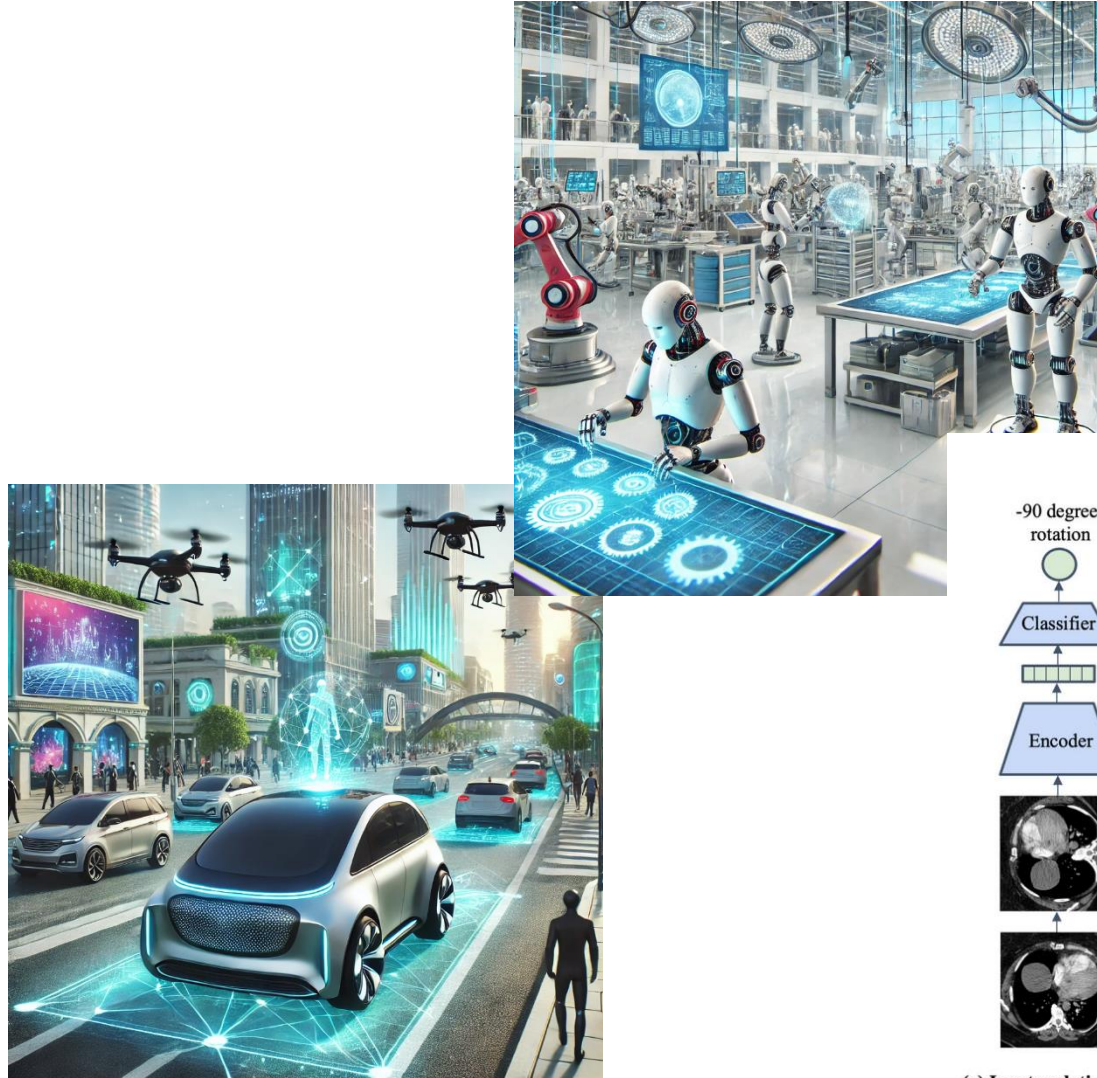
그 외

컴퓨터비전

자연어처리

• 그 외

자율주행, 로봇틱스, 의료 분야 등등... 수많은 도메인에서 딥러닝은 필수가 되었다.





감사합니다

