Multimodal LLMs for Medical Time Series

Michael Staniek

Department of Computational Linguistics
Heidelberg University

October 14, 2024

Overview

1. Organisation

2. Papers

Grading

- 30% Presentation
- 30% Participation
- 40% Report/Project

Presentation

Presenters work

- Read the paper, understand it, present it
- Either:
  - Prepare 2 questions for your own paper that the others have to answer
  - Prepare the answers to the questions that others send to you
- Direct the discussion at the end

Participants work

- Read the paper, understand it,
- Either:
  - Answer 2 questions for that paper
  - Send some questions to the corresponding papers
- Participate

Topic Distribution

- Please send me 3 paper preferences until the 20.10.2024 so that I can puzzle out who gets which paper
- Its not first come first served, so take your time. If more people want a paper, I start rolling in Python
- I planned with less than 14 people that want ECTS. If we are more, I will select more papers, or if you are feeling adventerous, you can send me paper proposals

Project/Term Paper

- Deadline: 31.03.2025 23:59:59:999
- But: You can already start now
- Project does not have to depend on your chosen paper, same with term paper
- If we have much fewer people participating, I'm also fine with second presentations.

The Task

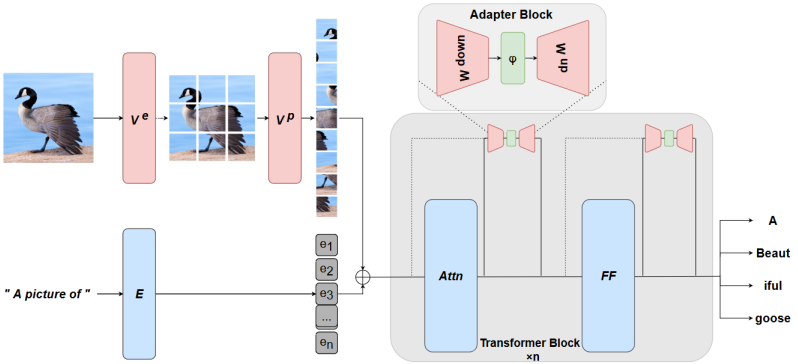


Figure: The general task that has to be solved

Multimodality

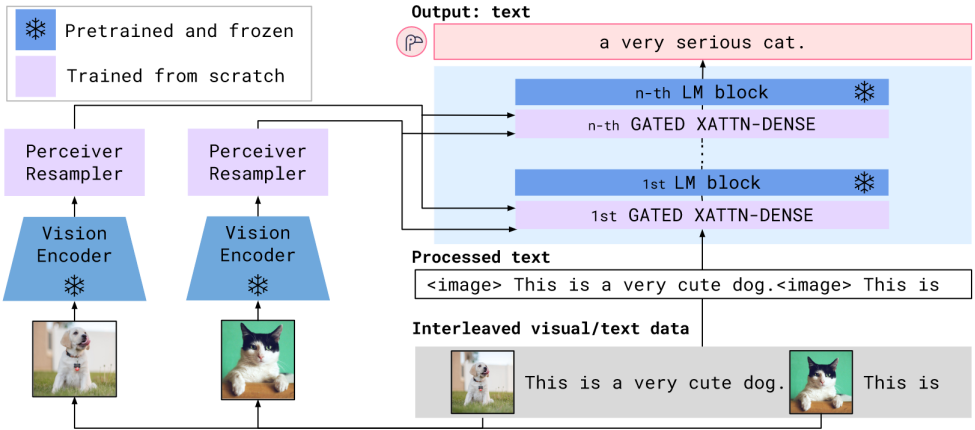
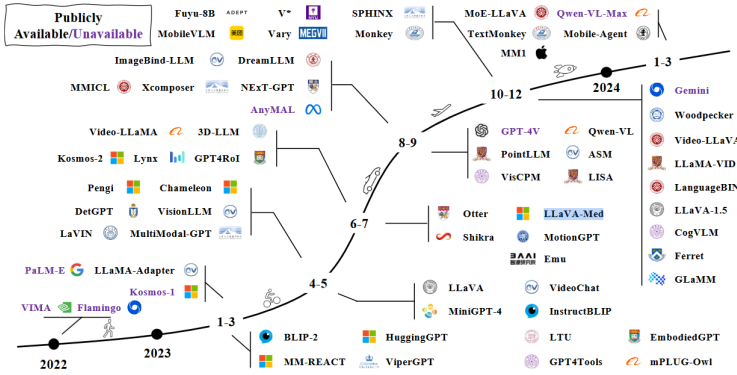


Figure: The general task that has to be solved

Multimodality is Hype



Very Hype

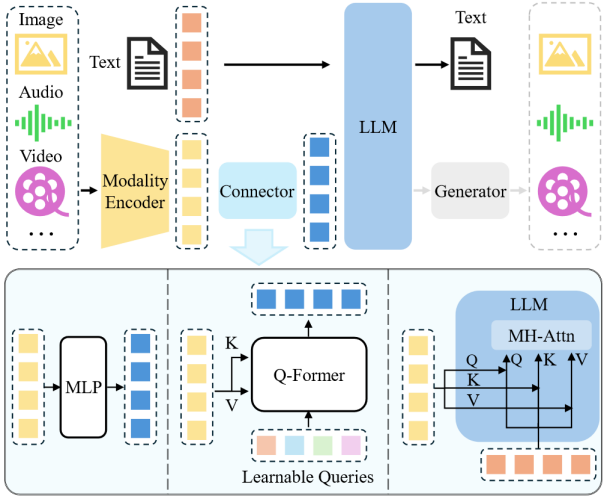


Image
Audio
Video
...
Text
Modality Encoder
Connector
LLM
Text
Generator
...
MLP
K
V
Q-Former
Learnable Queries
LLM
MH-Attn
Q
K
V
Q
K
V

Paperlist 1: Basics of Multimodality

1. MAGMA – Multimodal Augmentation of Generative Models through Adapter-based Finetuning
2. BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models
3. Flamingo: a Visual Language Model for Few-Shot Learning

Paperlist 2: Basics of Medical Data

1. Self-supervised transformer for sparse and irregularly sampled multivariate clinical time-series
2. Research on Multimodal Fusion of Temporal Electronic Medical Records
3. PromptEHR: Conditional Electronic Healthcare Records Generation with Prompt Learning

Paperlist 3: (Medical) Time Series LLMs

1. A Multimodal Transformer: Fusing Clinical Notes with Structured EHR Data for Interpretable In-Hospital Mortality Prediction
2. Time Series as Images: Vision Transformer for Irregularly Sampled Time Series
3. Deep multi-modal intermediate fusion of clinical record and time series data in mortality prediction
4. FuseMoE: Mixture-of-Experts Transformers for Fleximodal Fusion
5. Learning Missing Modal Electronic Health Records with Unified Multi-modal Data Embedding and Modality-Aware Attention
6. Integrating Physiological Time Series and Clinical Notes with Transformer for Early Prediction of Sepsis
7. Multimodal Pretraining of Medical Time Series and Notes
8. Research on Multimodal Fusion of Temporal Electronic Medical Records

Paperlist 4: Additional Papers

1. Learning to write notes in electronic health records
2. The shaky foundations of large language models and foundation models for electronic health records.
3. Improving Medical Predictions by Irregular Multimodal Electronic Health Records Modeling
4. Time Series as Images: Vision Transformer for Irregularly Sampled Time Series
5. Hierarchical Pretraining on Multimodal Electronic Health Records
6. Subtle variation in sepsis-III definitions markedly influences predictive performance within and across methods

Learning to write notes in electronic health records

The End