Image Segmentation

Introduction to Segment Anything

Jan Theiß

December 11, 2023

Outline

What are Transfromers?

- We think about the concept of transformers most of the time in Natural Language Processing(NLP)-Tasks
- Transformers can modify (transform) the input
- A very basic transformer could be a translator

Attention Is All You Need

Target:

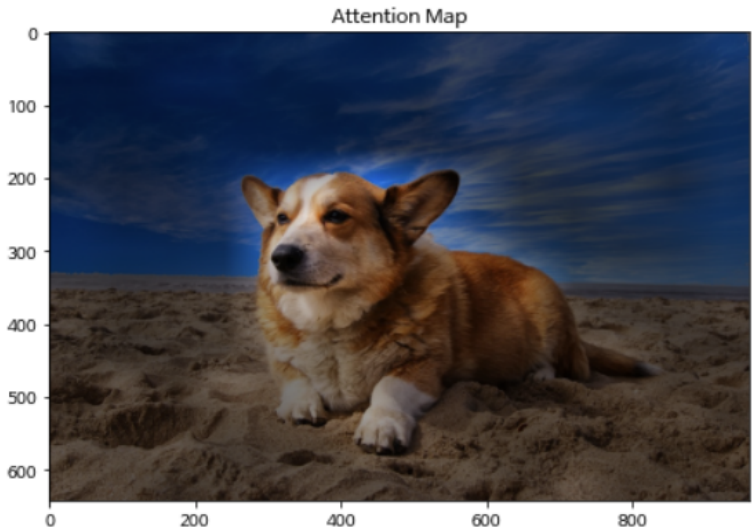
- Semantic Translation
- from word-for-word to sense-for-sense

$\Longrightarrow$ pairwise inner product of every item of the input set (i.e. words in a sentence)

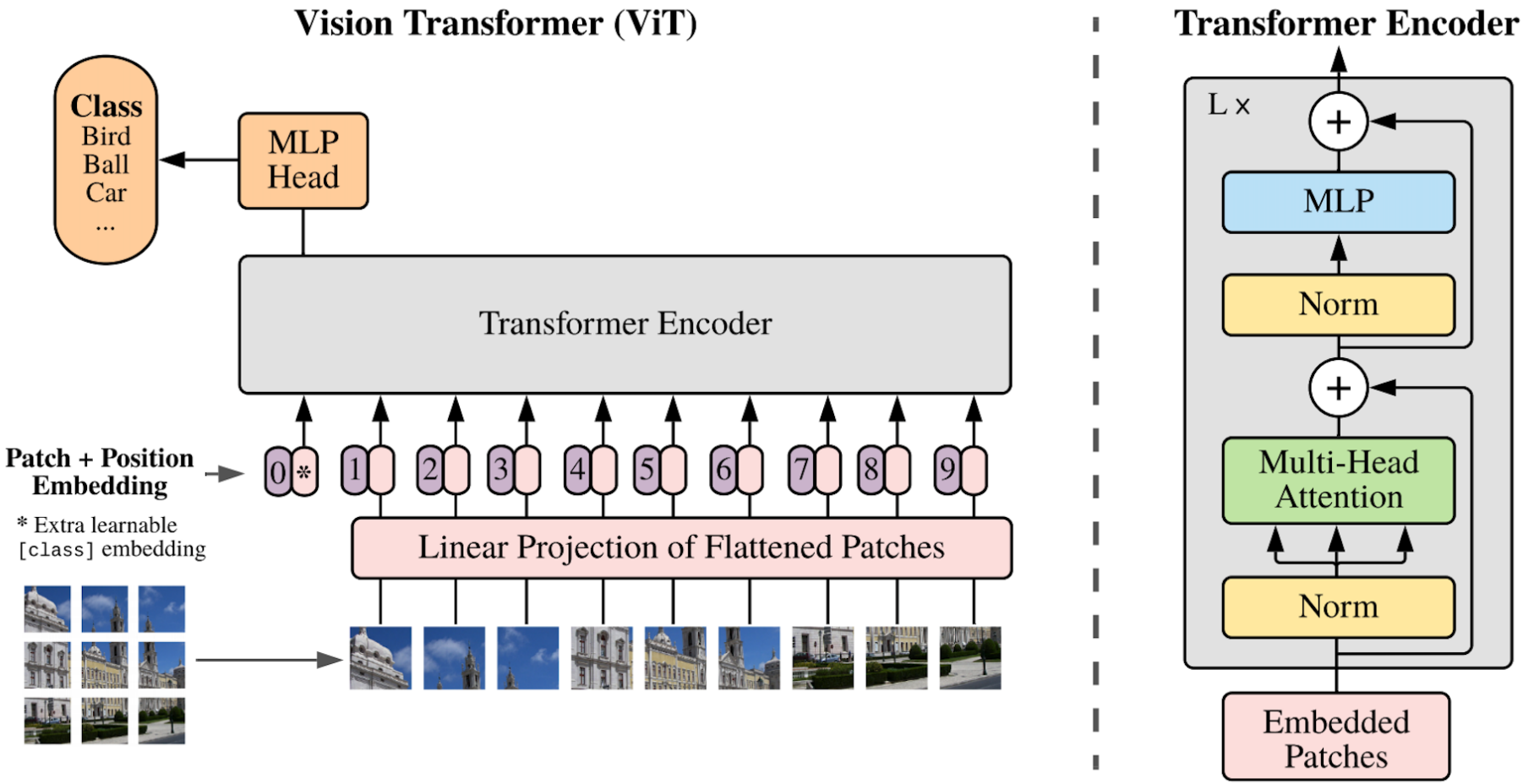
Problem when dealing with images as input:

250x250 $\Rightarrow$ 62.500 px $\Rightarrow$ 3.906.250.000 calculations

Attention Is All You Need



An Image Is Worth 16x16 Words



What is Segment Anything?

Video Demonstration

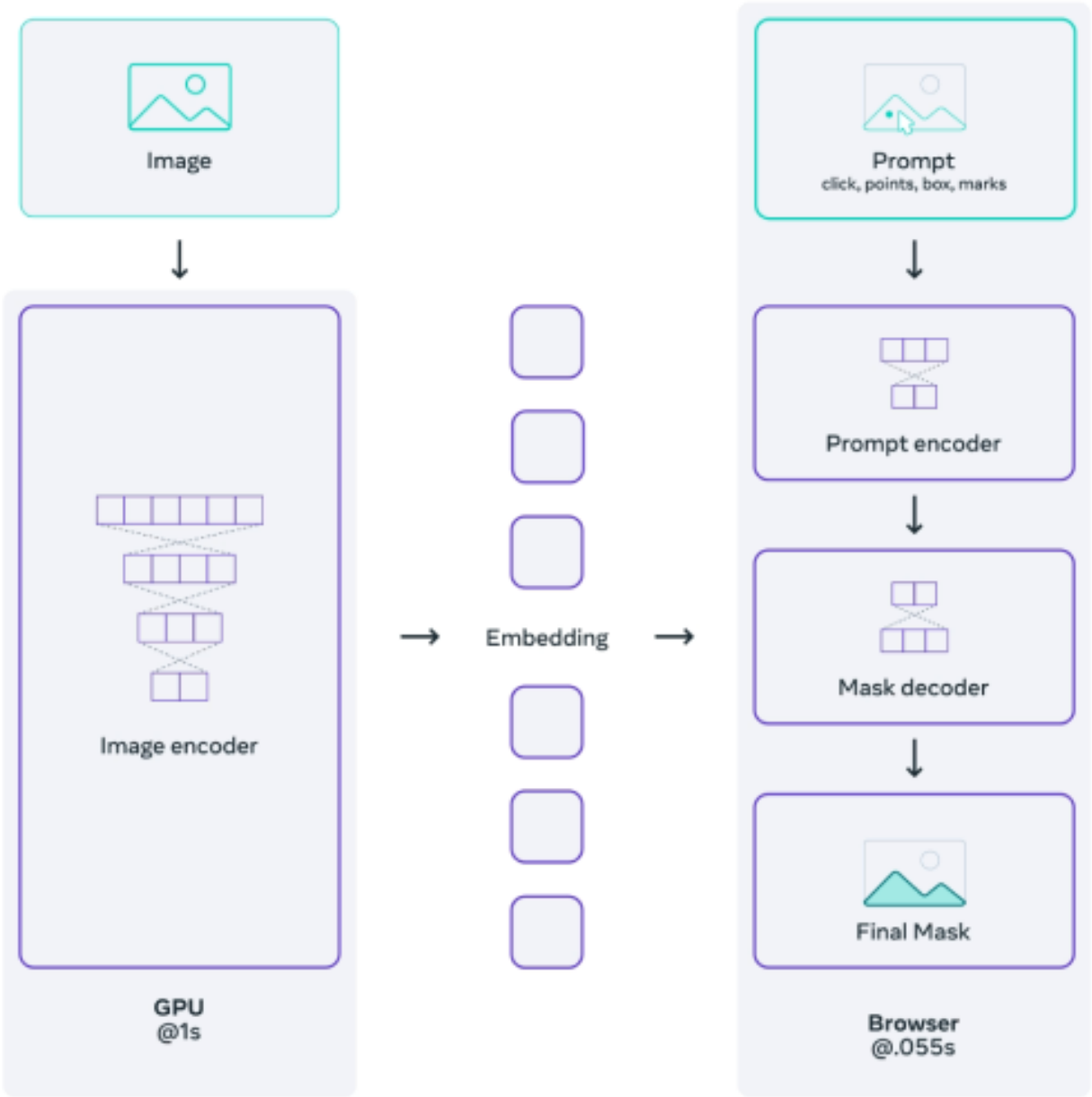
What is Segment Anything?

Definition And Key Aspects

Zero-Shot Large Language Model (LLM) developed by Meta AI that can segment objects in images

1. Tasks
   - How to interact with the model and what inputs are available
2. Model
   - Segment Anything Model (SAM)
3. Data
   - SA-1B

What does SAM look like?



SA-1B Dataset

- 11M images
- 1.1B masks (99.1 % automatically collected)

Software Demo

DEMO
https://segment-anything.com/demo

Implementing SAM

1. go onto github https://github.com/facebookresearch/segment-anything/tree/main
2. find an image of your choice
3. use opencv (or matplotlib.pyplot) to load the image
4. initiate SAM and a suitable predictor with the image
5. preview the masks