

Presentación

Segment Anything

Meta AI Research FAIR



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE

Contexto y objetivo

NLP – Natural Language Processing

- Grandes dataset
- Foundations models
- Prompting



Zero-shot y few-shot generalización

Computer Vision – Segmentación de imágenes

- Ningún gran dataset con masks



OBJETIVO : Desarrollar un “foundation model” para la segmentación de imágenes

SAM – Funcionamiento

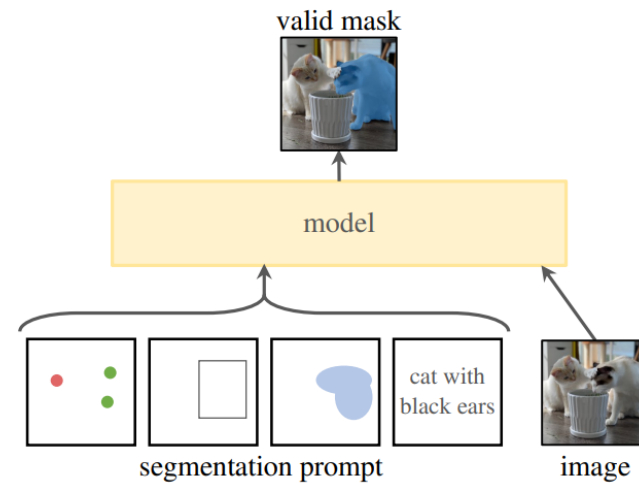
Cualquier prompt



Mask de segmentación válido



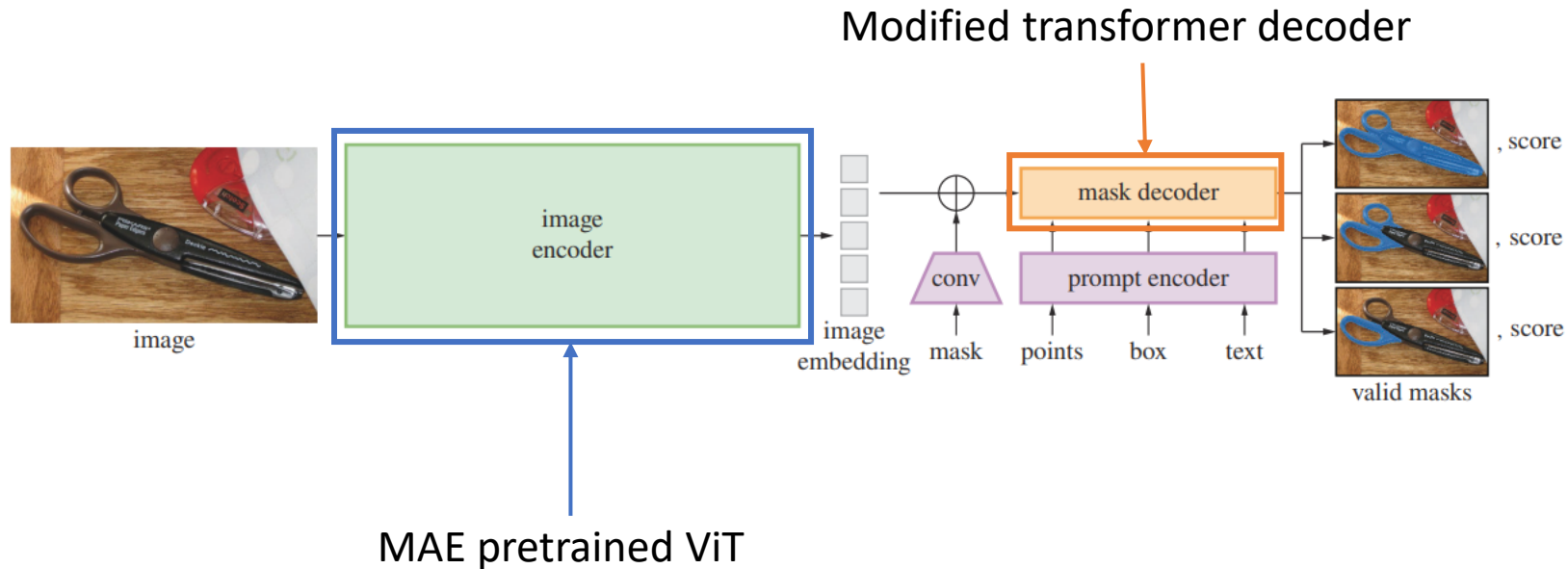
Ambiguamente consciente



(a) **Task:** promptable segmentation

SAM – Modelo

Segment Anything Model (SAM) overview



Data engine – SA-1B



3 fases

①

Asistido-manual :

Primero se entrena con dataset público
Luego con máscaras recién anotadas
Encoder escalado de ViT-B a ViT-H

4,3M masks

②

Semiautomático :

Primero, detección automática de masks de confianza
Luego, profesionales anotan objetos adicionales

5,9 M masks

③

Totalmente automatizado :

Anotación totalmente automática
Predicción IoU para los masks seguros
Selección de masks estables

Data Set



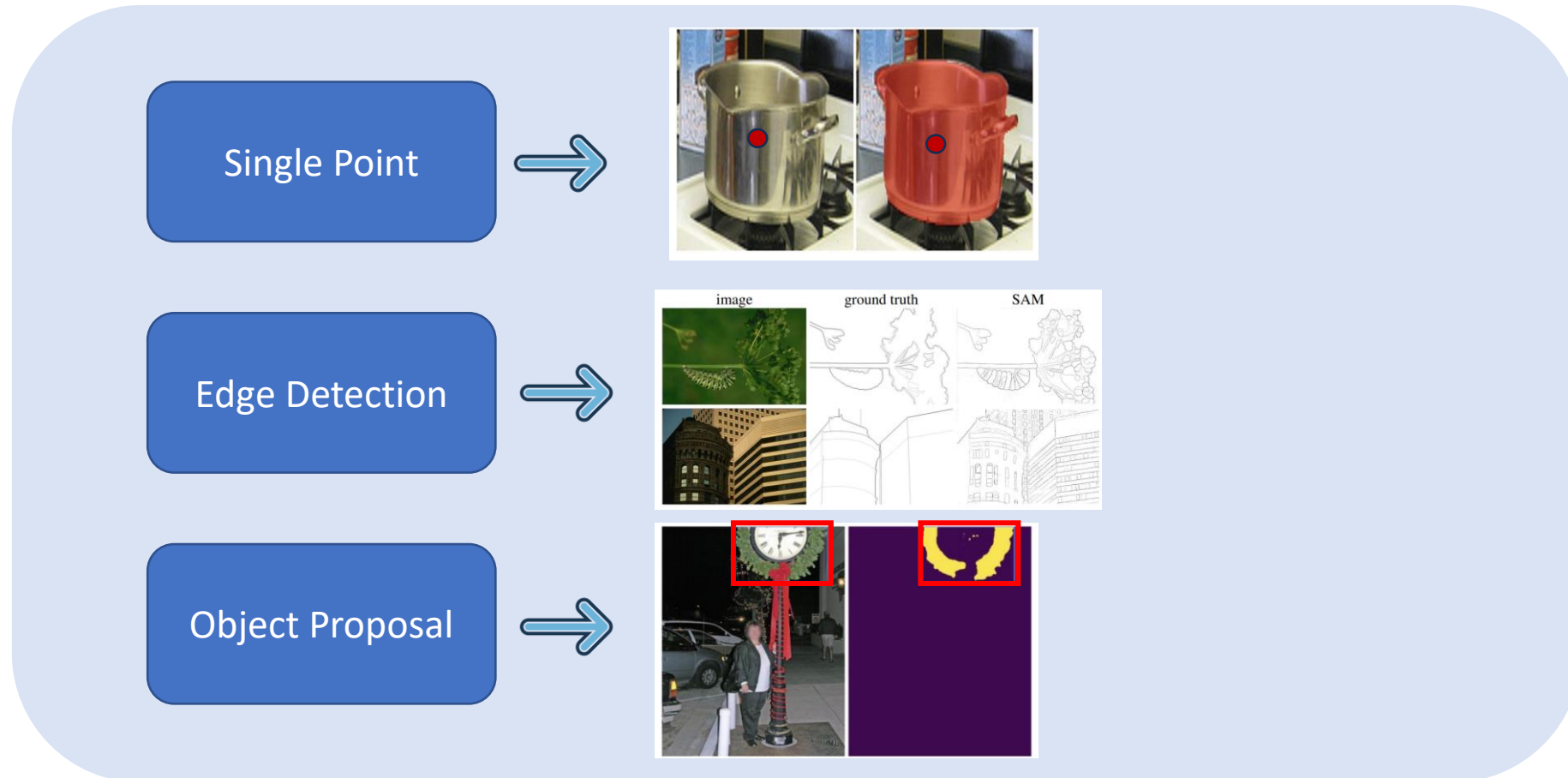
High Quality Image (3300×4950 pixels en promedio)

1,1B Masks

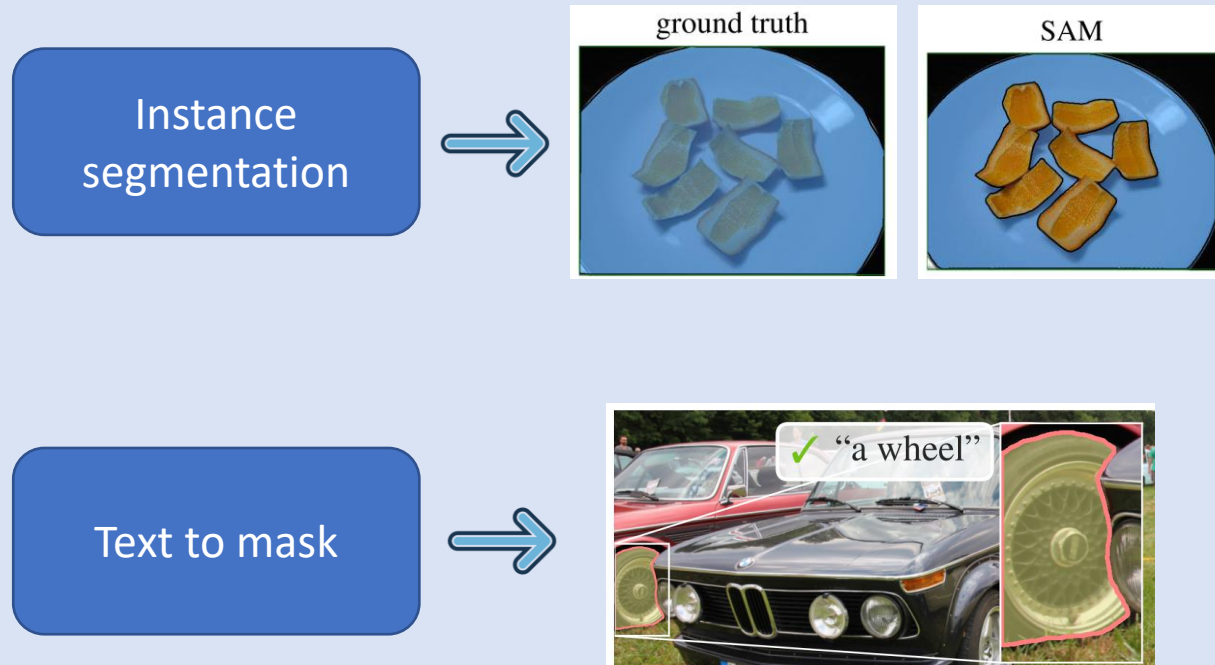
99,1 % generado
automáticamente

Zero-Shot Transfer Experiment

→ Experiment on 23 datasets



Zero-Shot Transfer Experiment



Zero-Shot Transfer Experiment – Resultados

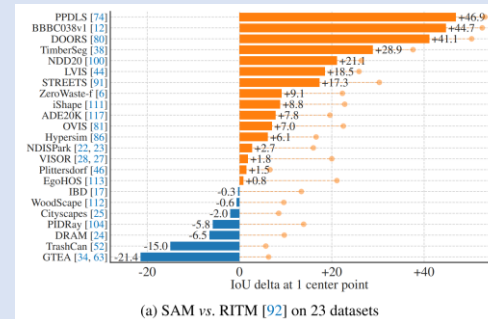
Single Point



Edge Detection



Object Proposal



SAM vs RITM

➔ Mejor en situaciones ambiguas

- Predice más bordes que la verdad
- Insuficiente en comparación con state of the art
- Supera otros métodos DL

method	all	mask AR@1000					
		small	med.	large	freq.	com.	rare
ViTDet-H [62]	63.0	51.7	80.8	87.0	63.1	63.3	58.3
zero-shot transfer methods:							
SAM – single out.	54.9	42.8	76.7	74.4	54.7	59.8	62.0
SAM	59.3	45.5	81.6	86.9	59.1	63.9	65.8

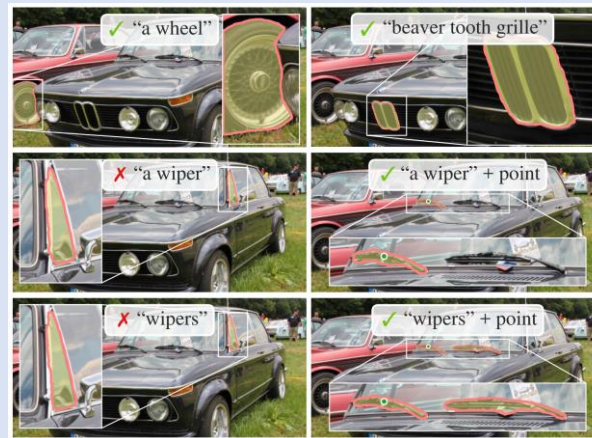
Zero-Shot Transfer Experiment – Resultados

Instance
segmentation



- Detrás de VitDet
- En un estudio humano, SAM supera a VitDet

Text to mask



Conclusión

Rendimiento competitivo

Limitaciones

Un análisis responsable interesante

Base de datos y modelo SAM públicos



Gracias por su atención



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE