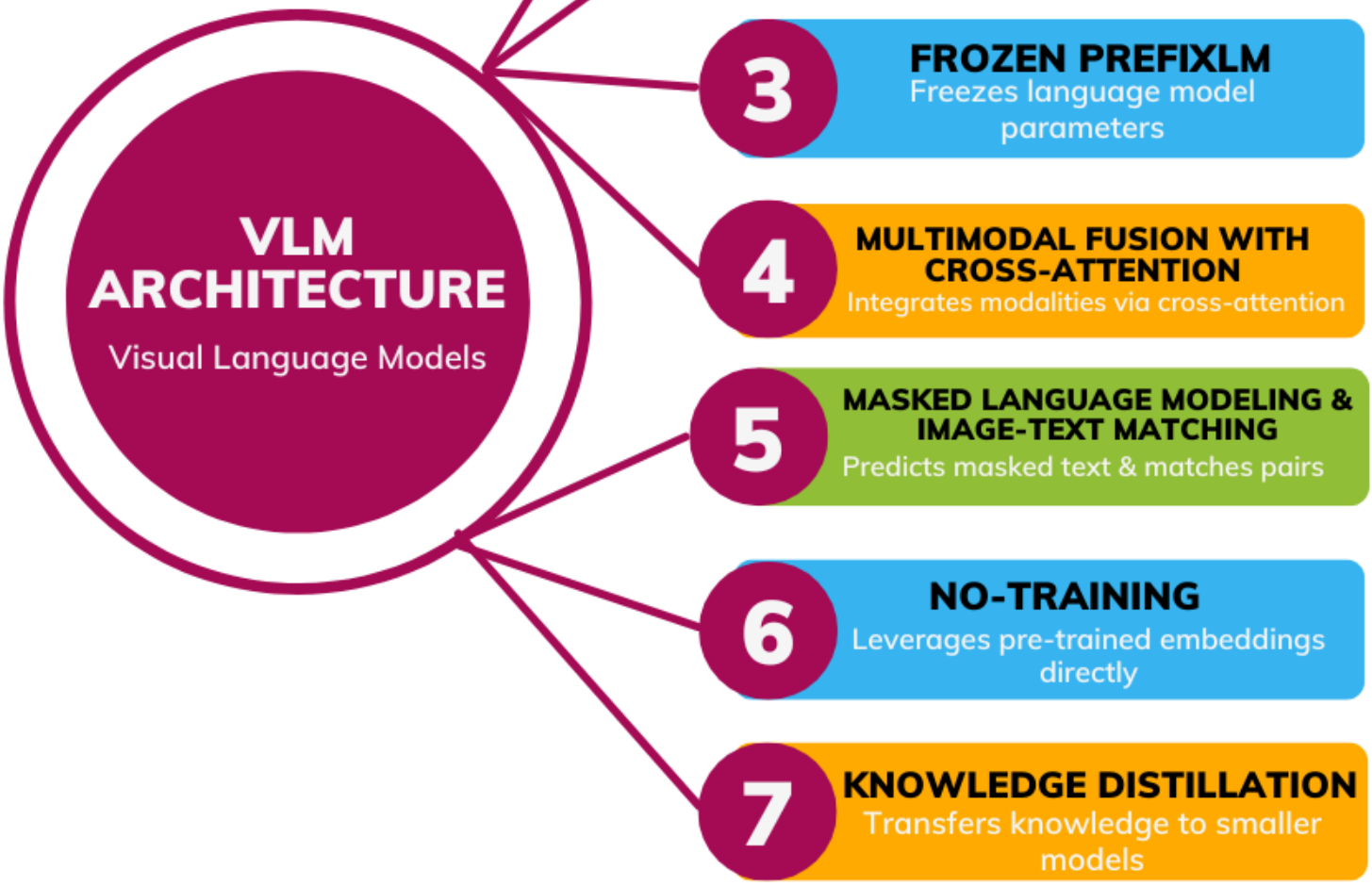
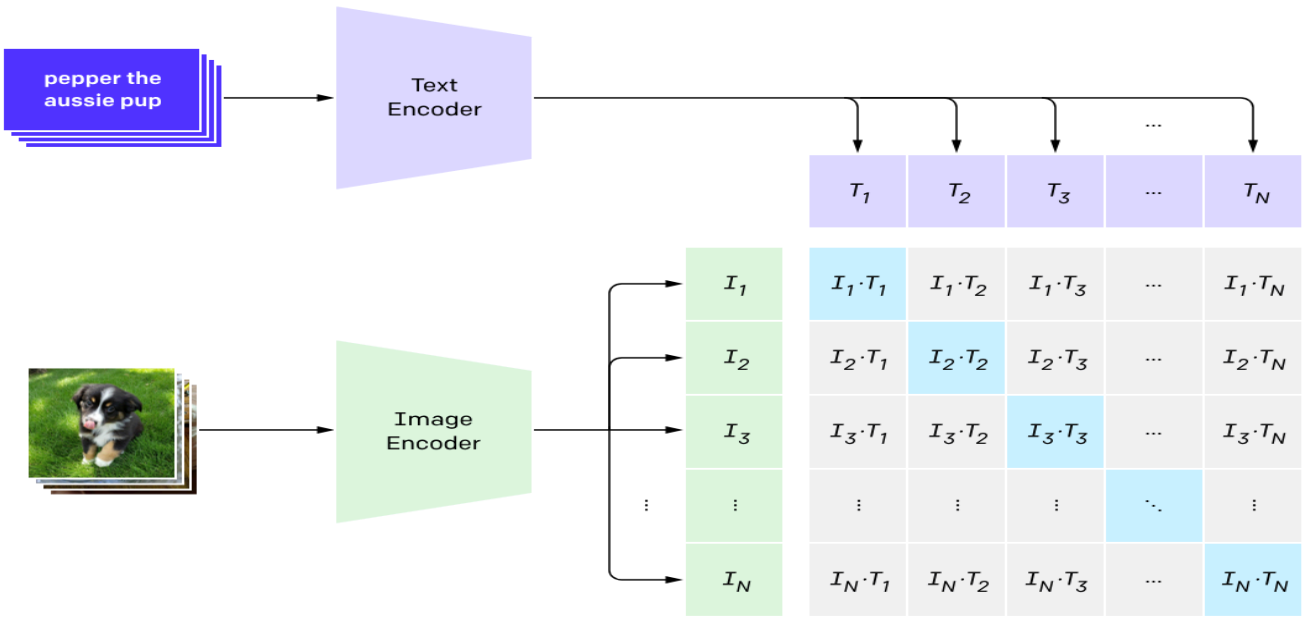
# VLM ARCHITECTURE



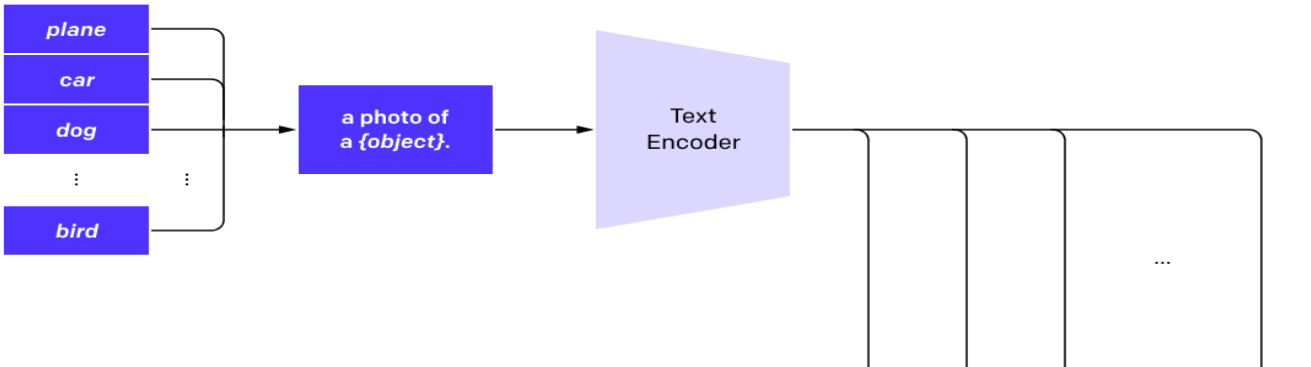
# Contrastive Learning

## CLIP Architecture

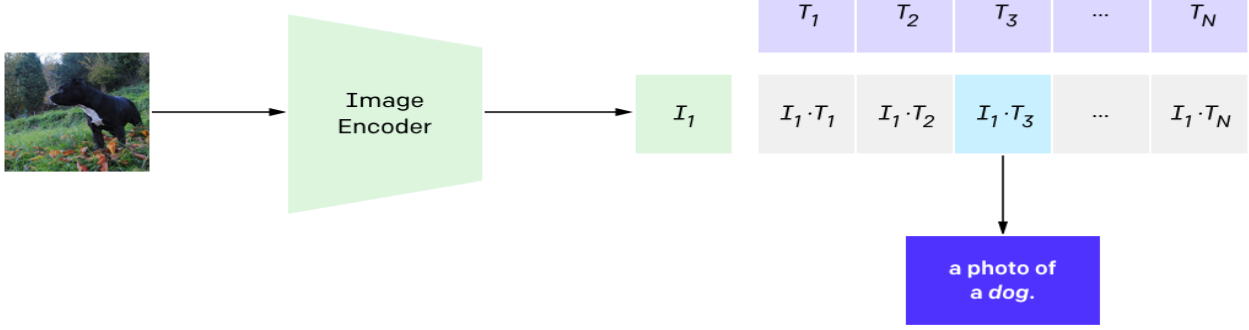
### 1. Contrastive pre-training




### 2. Create dataset classifier from label text

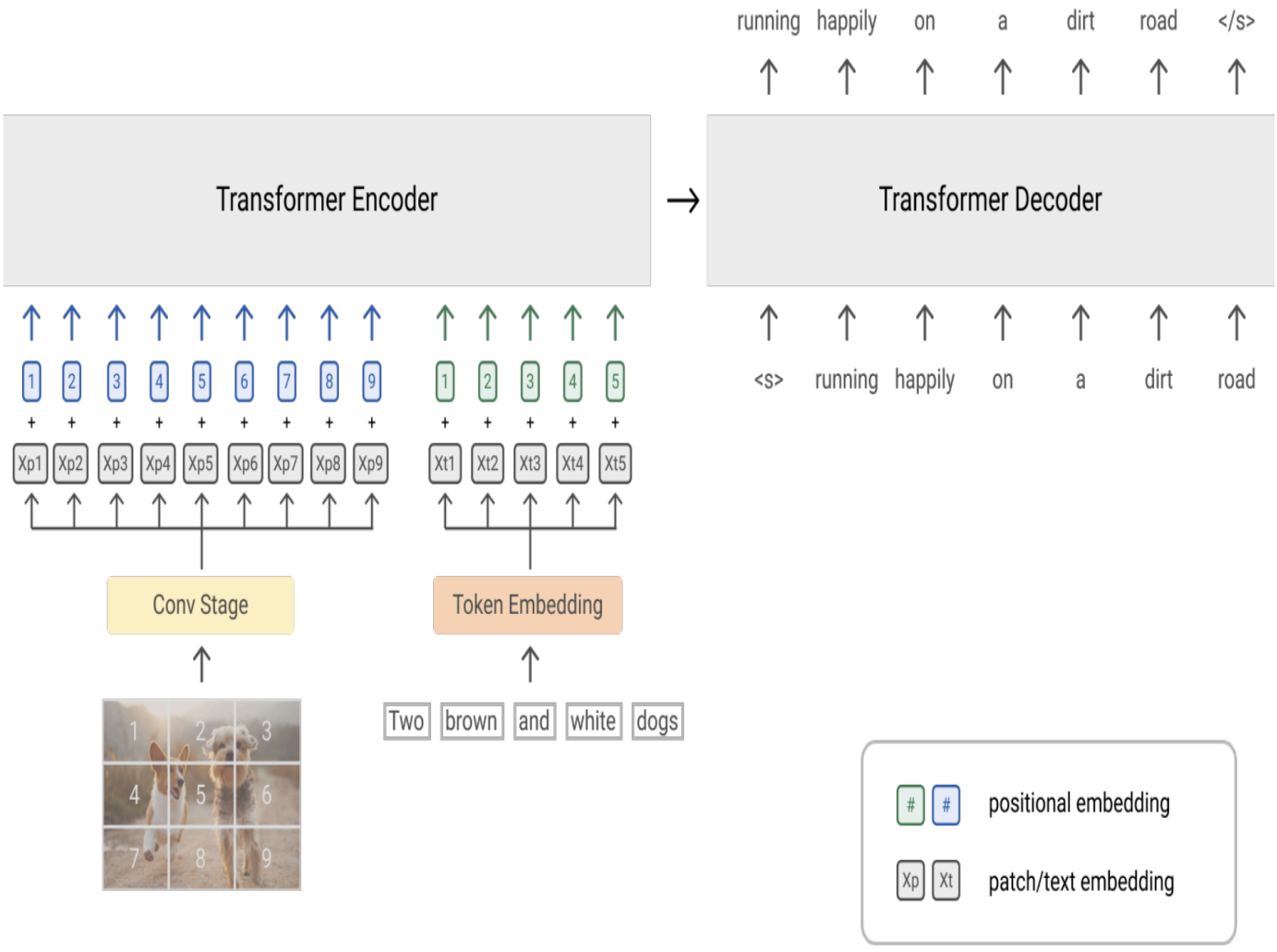



### 3. Use for zero-shot prediction



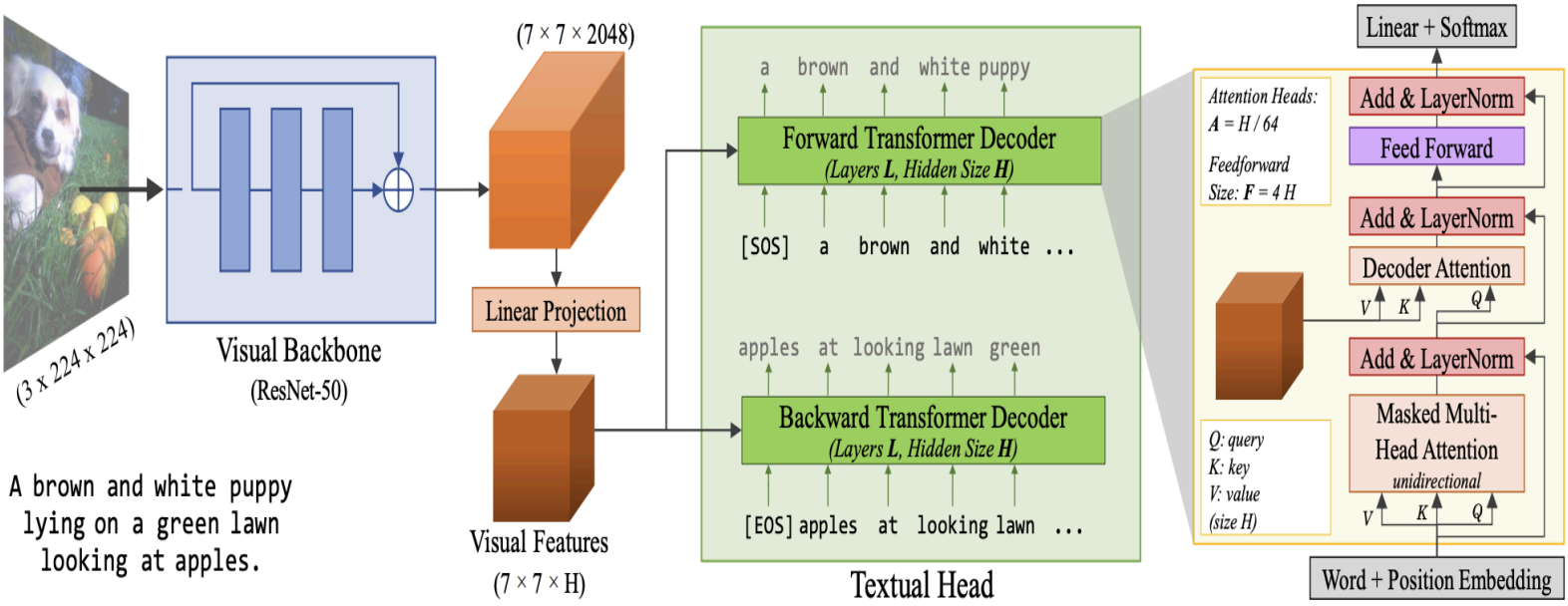
# PrefixLM

## SimVLM Architecture



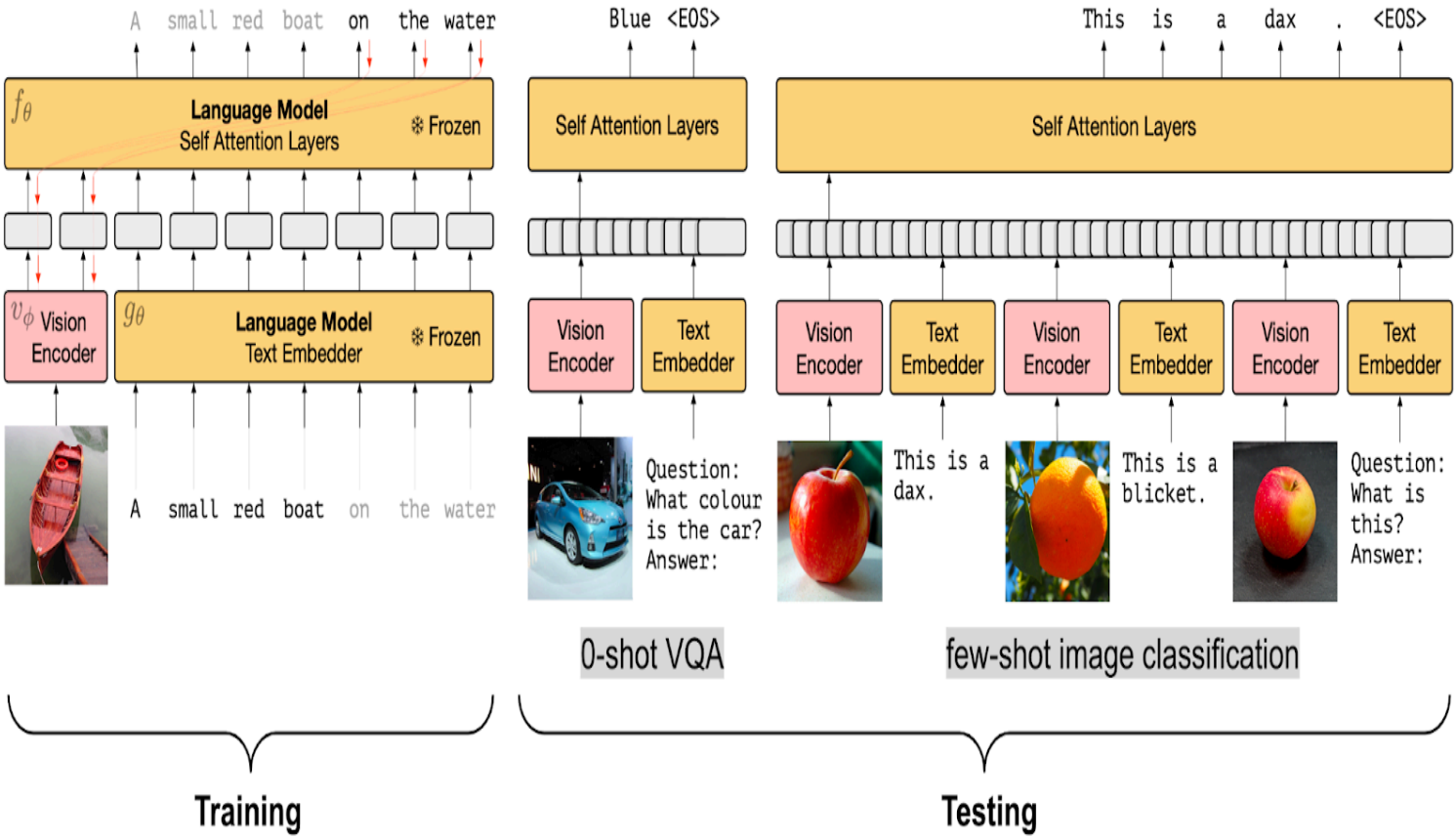
# PrefixLM

## VirTex Architecture



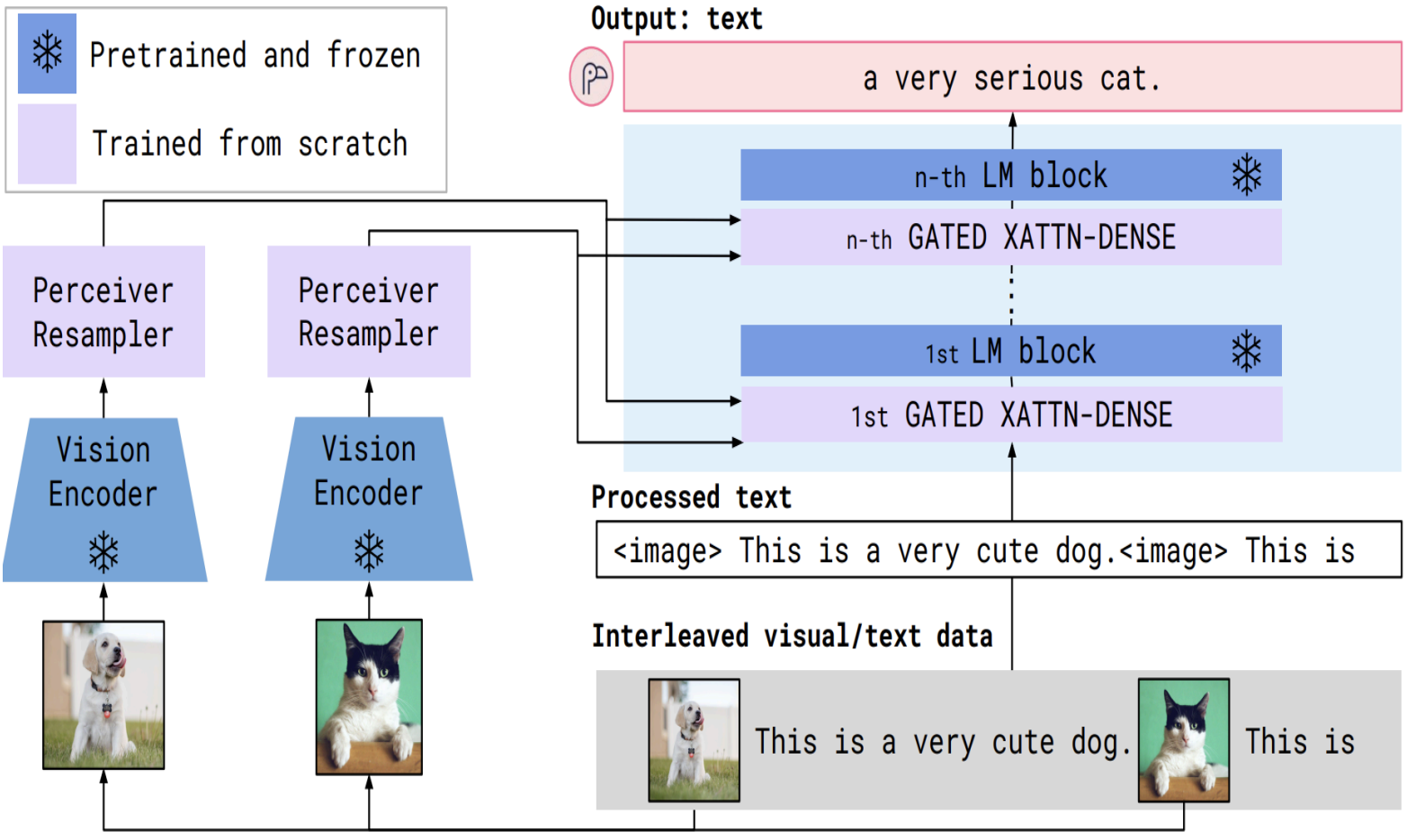
# Frozen PrefixLM

## Frozen Architecture



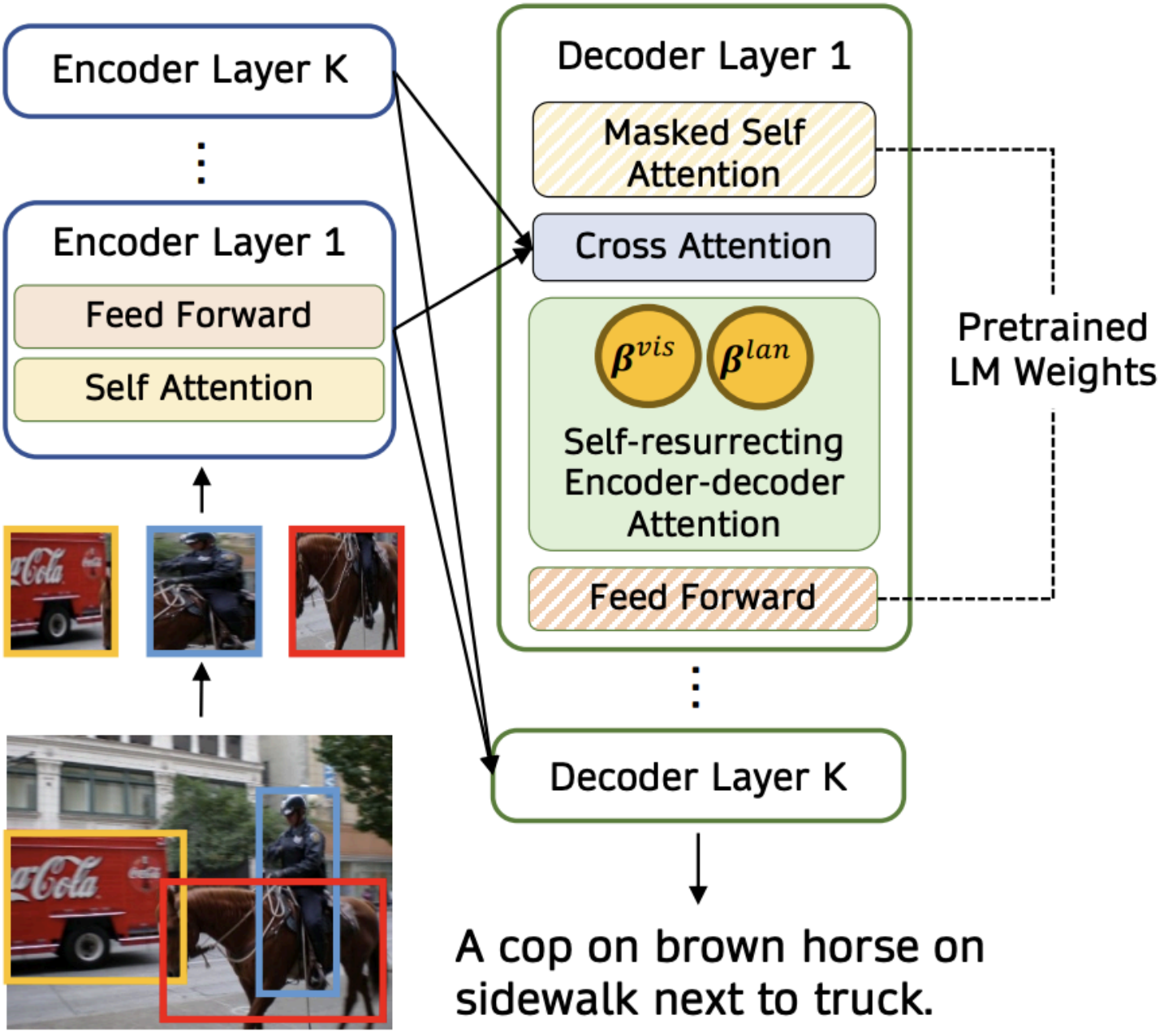
# Frozen PrefixLM

## Flamingo Architecture



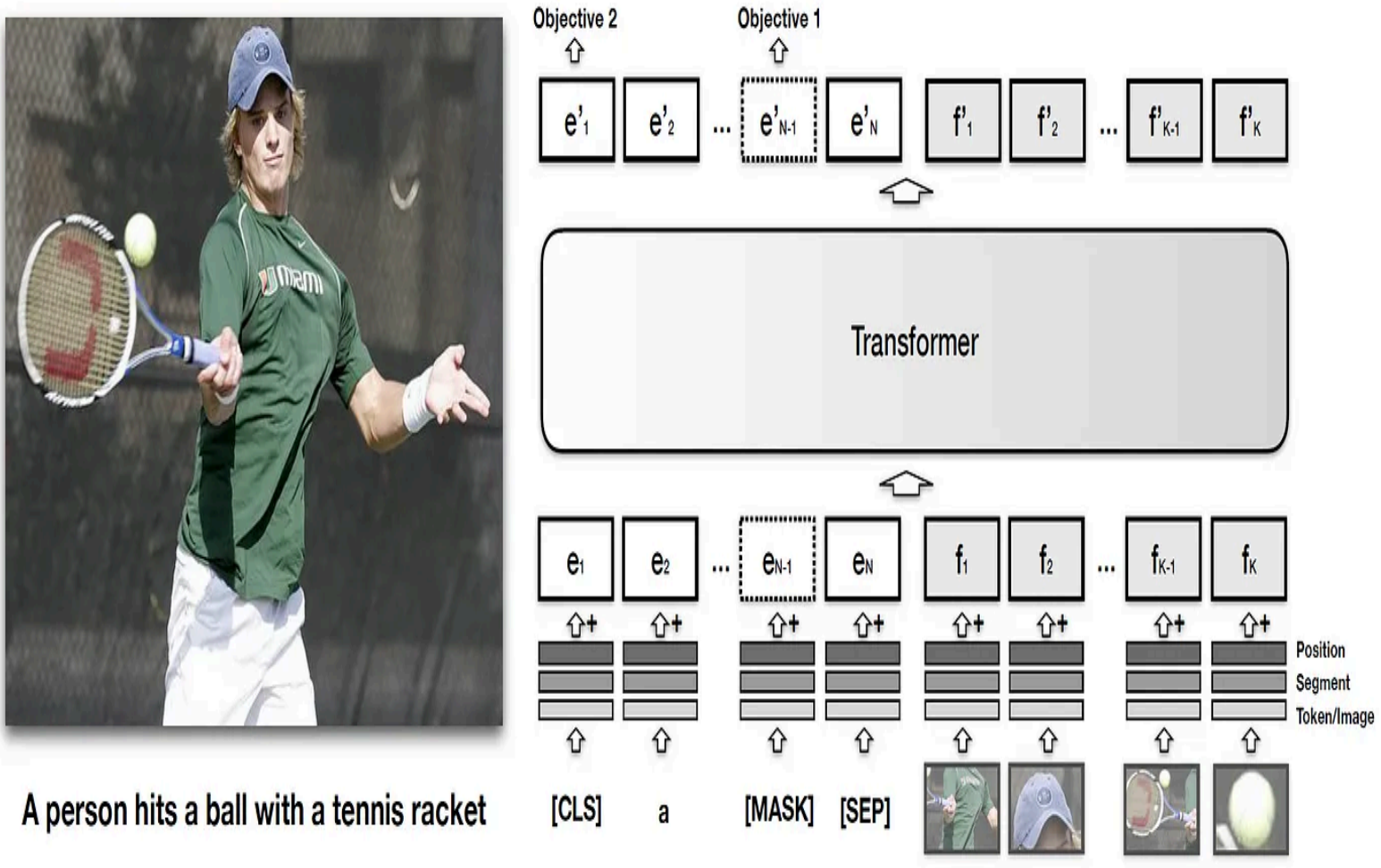
# Multimodal Fusing with Cross-Attention

## VisualGPT Architecture



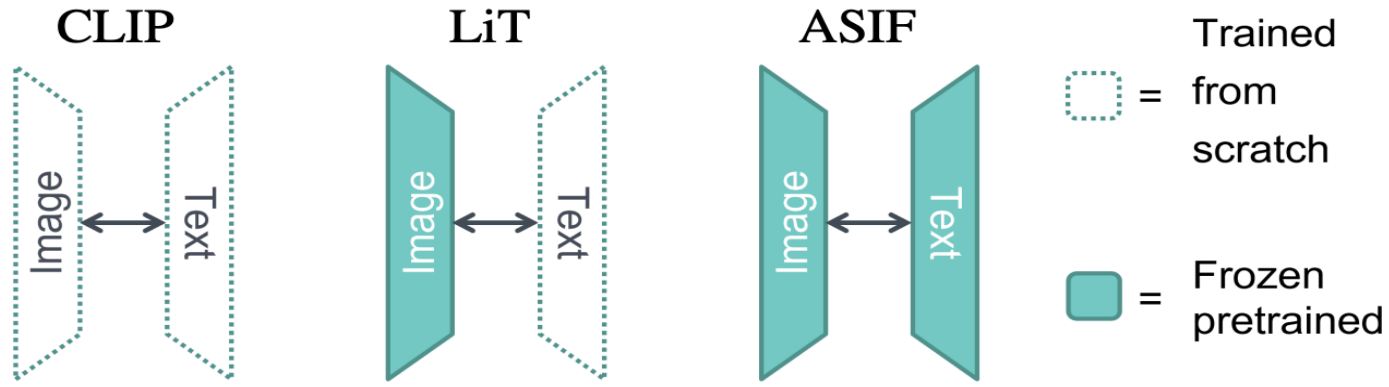
# Masked-language Modeling (MLM) and Image-Text Matching (ITM)

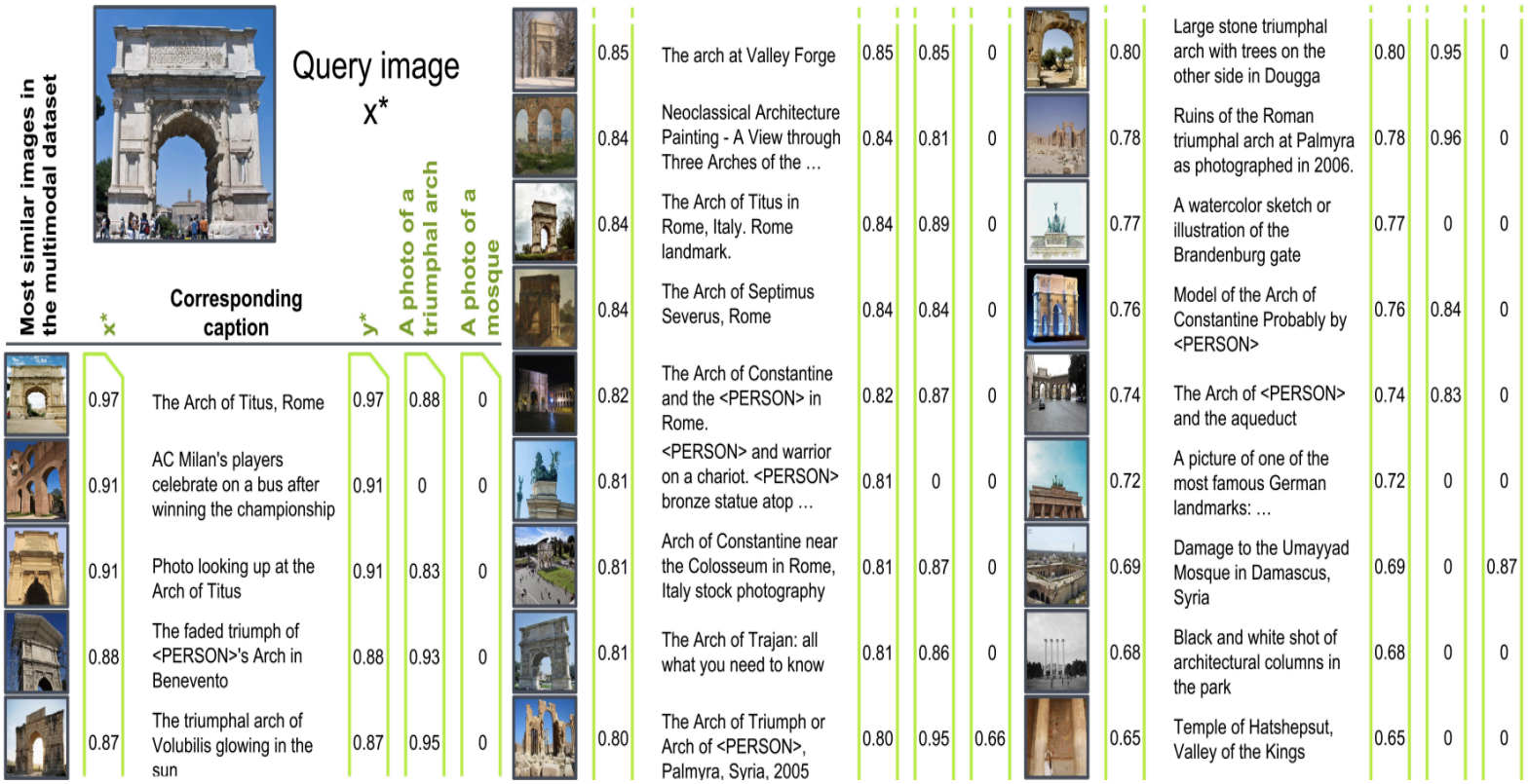
## VisualBERT Architecture



# No Training

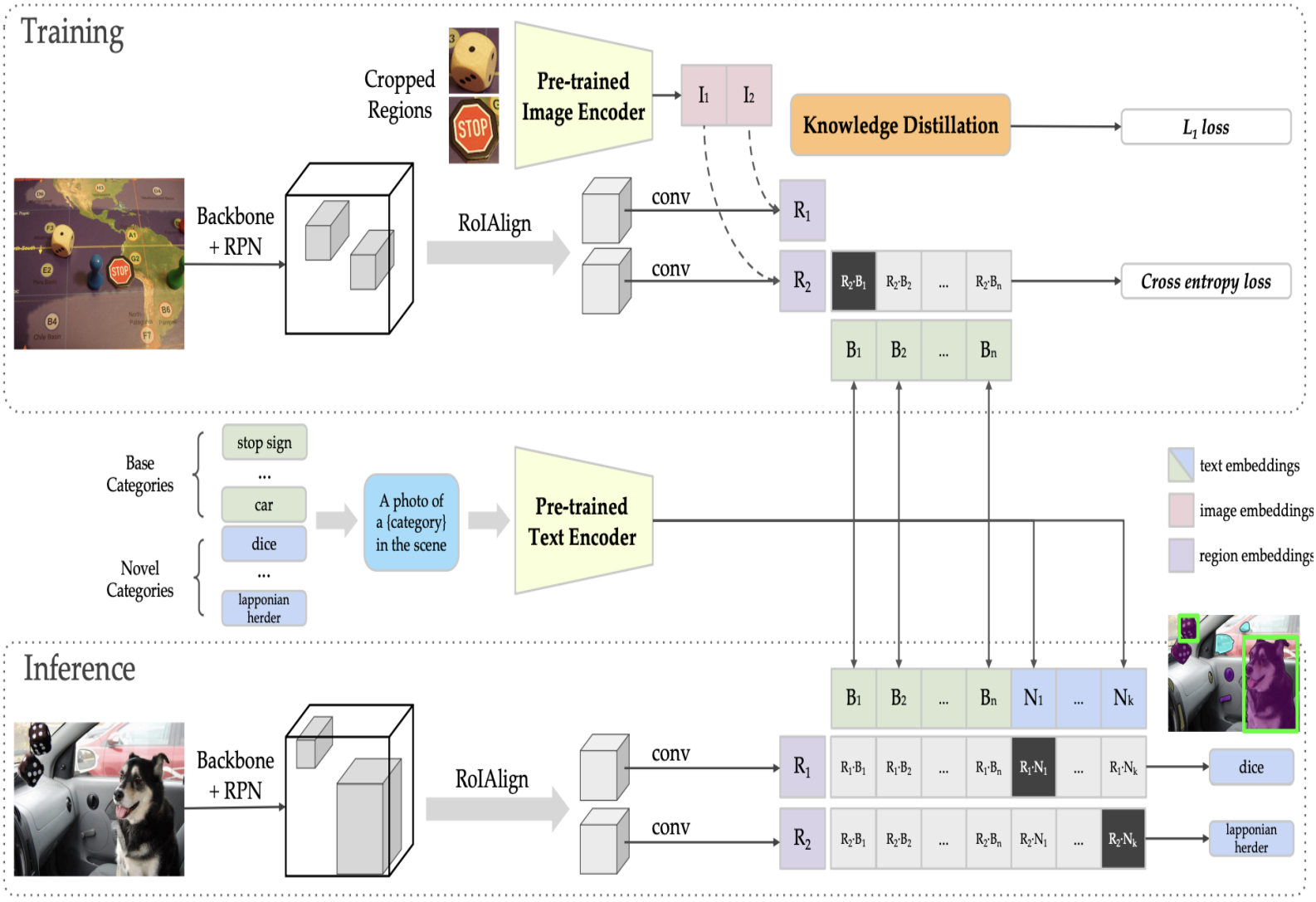
## ASIF Prediction Strategy



# Knowledge Distillation

## ViLD Architecture



**Download Complete Guide ⬇️**

https://github.com/Abonia1/VLM-Architecture