Week 1

at the ucf reu in computer vision

Charlie Levy

Introduction







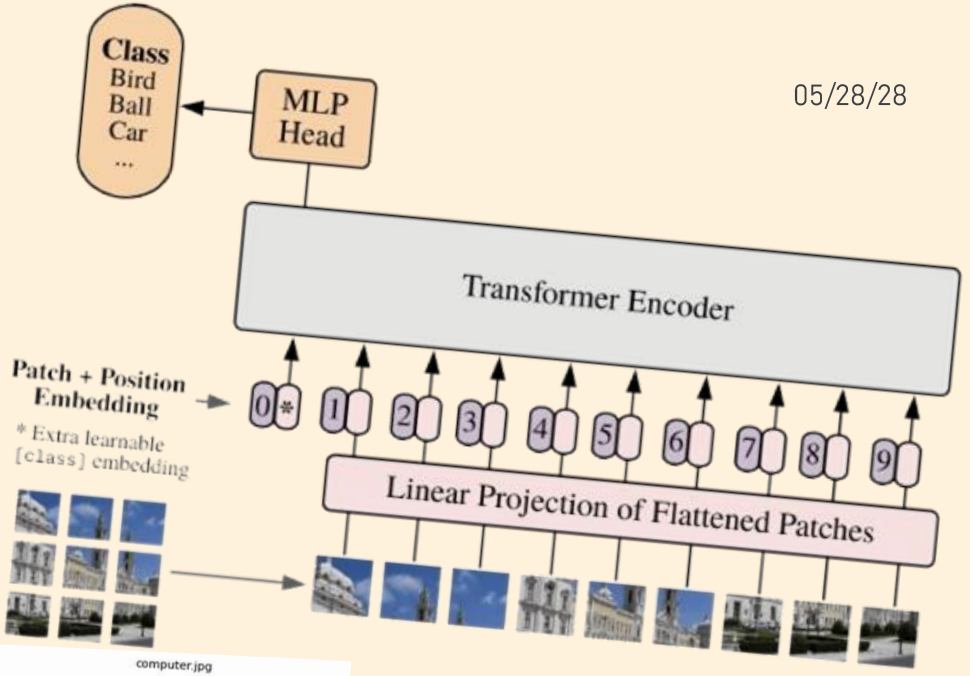
Classes I've Taken:
Combinatoric Structures
Computer Systems
Analysis of Algorithms
Geometric Algorithms
Probability in Computing
Database Systems

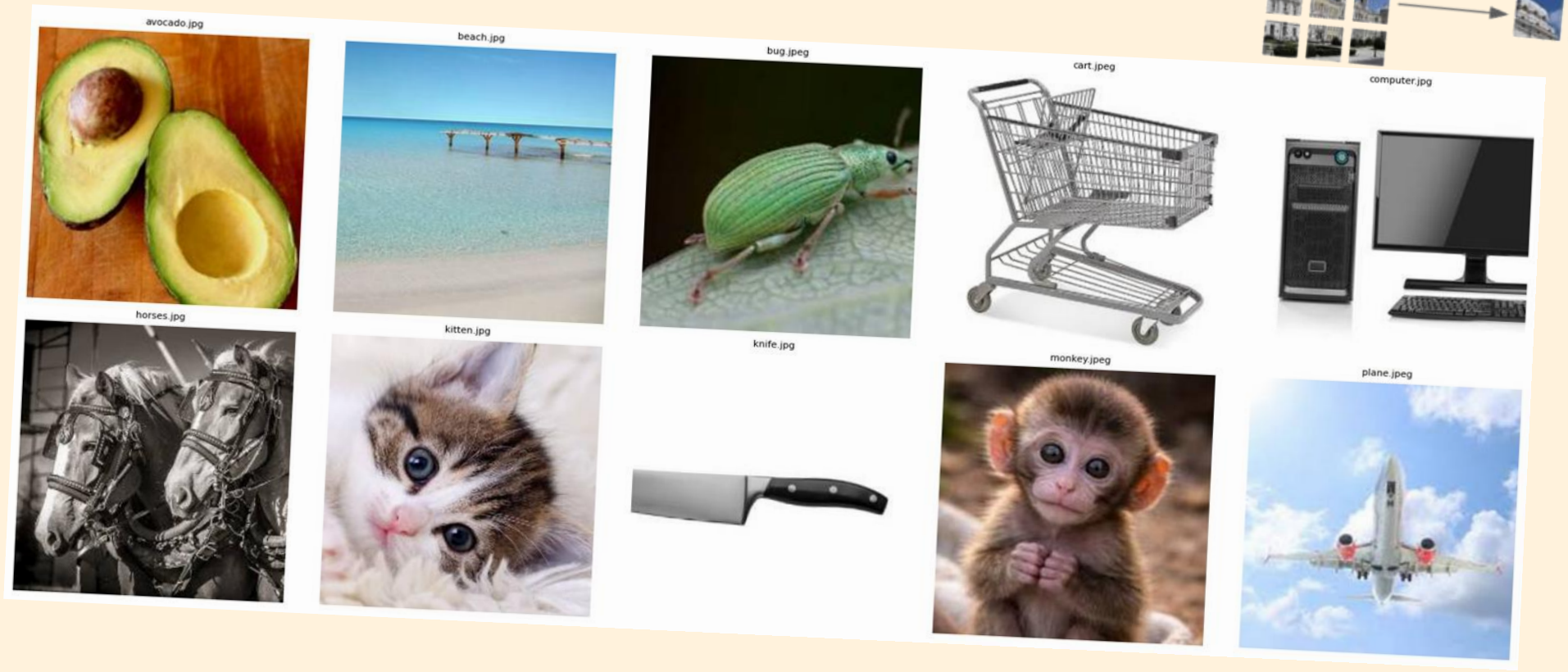


Charlie Levy

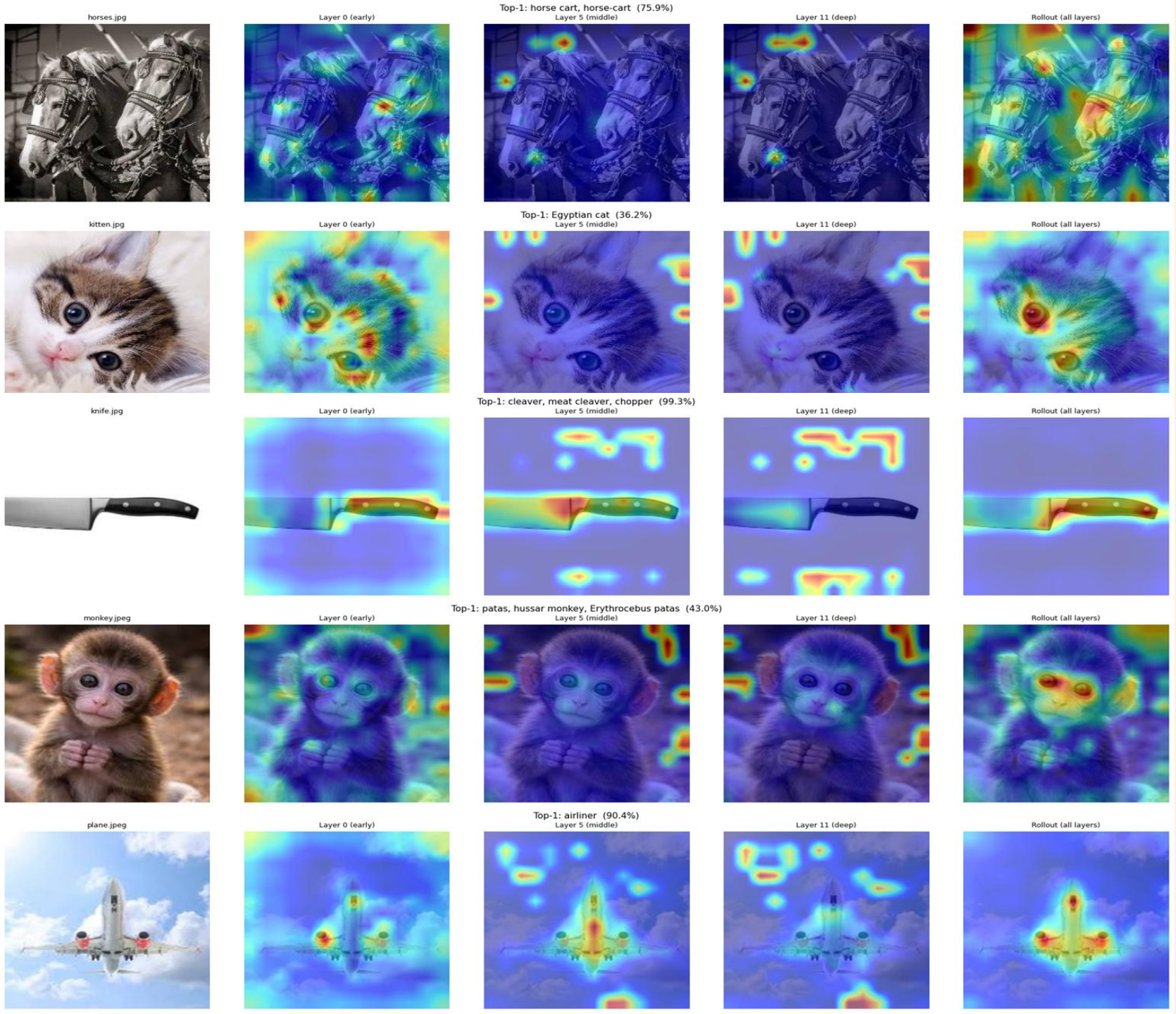
Assignment 1







ATTENTION MAPS



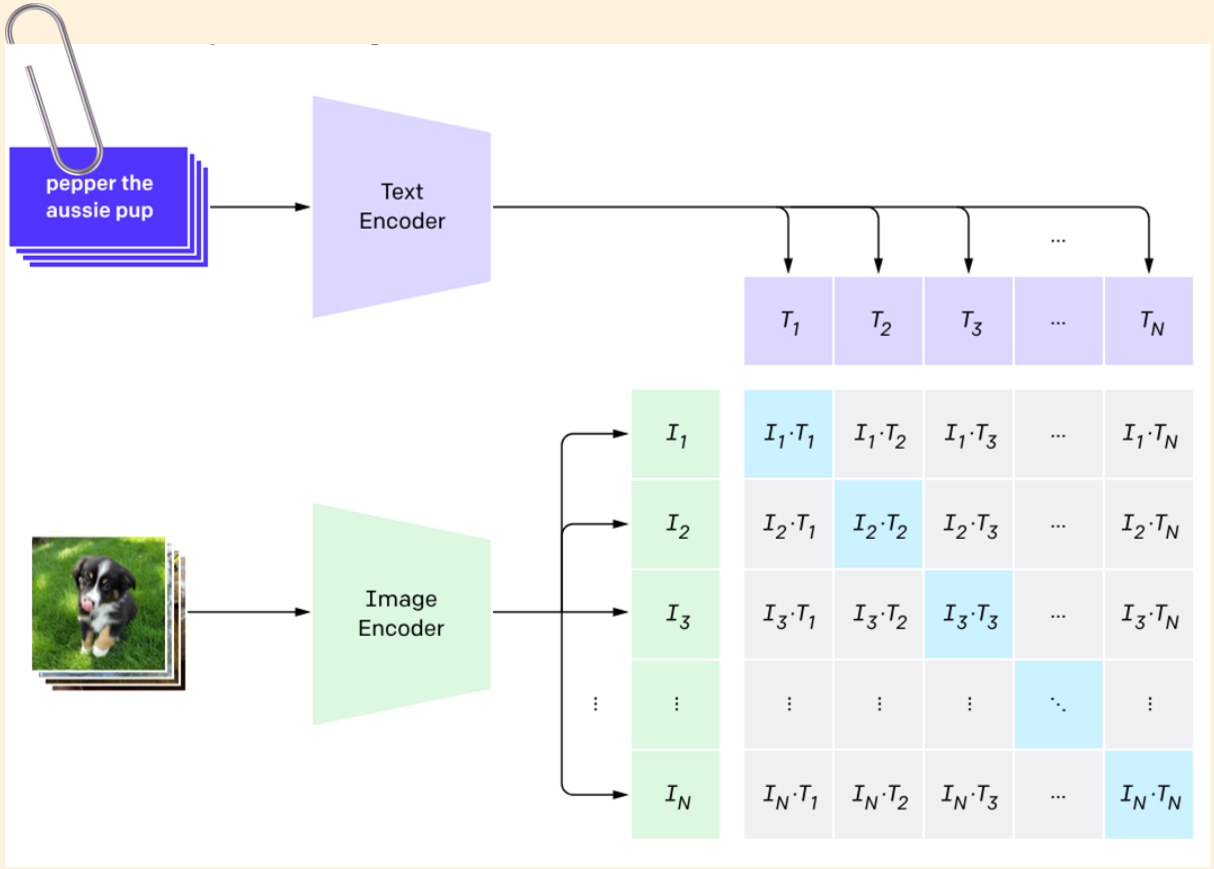
ATTENTION MAPS

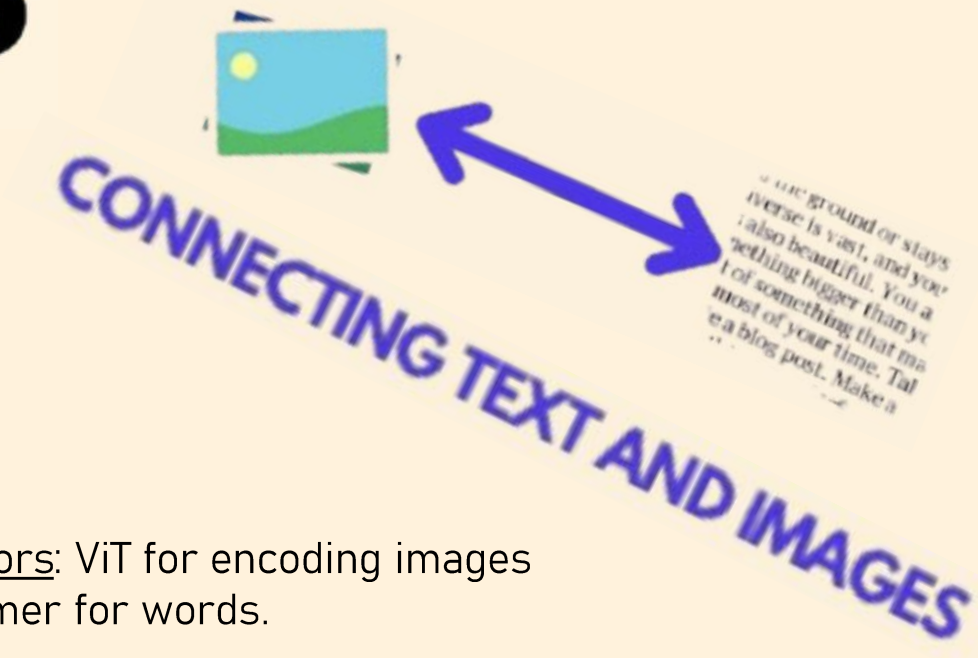
Assignment 2



OpenAI CLIP

(Contrastive Language-Image Pre-training) ==

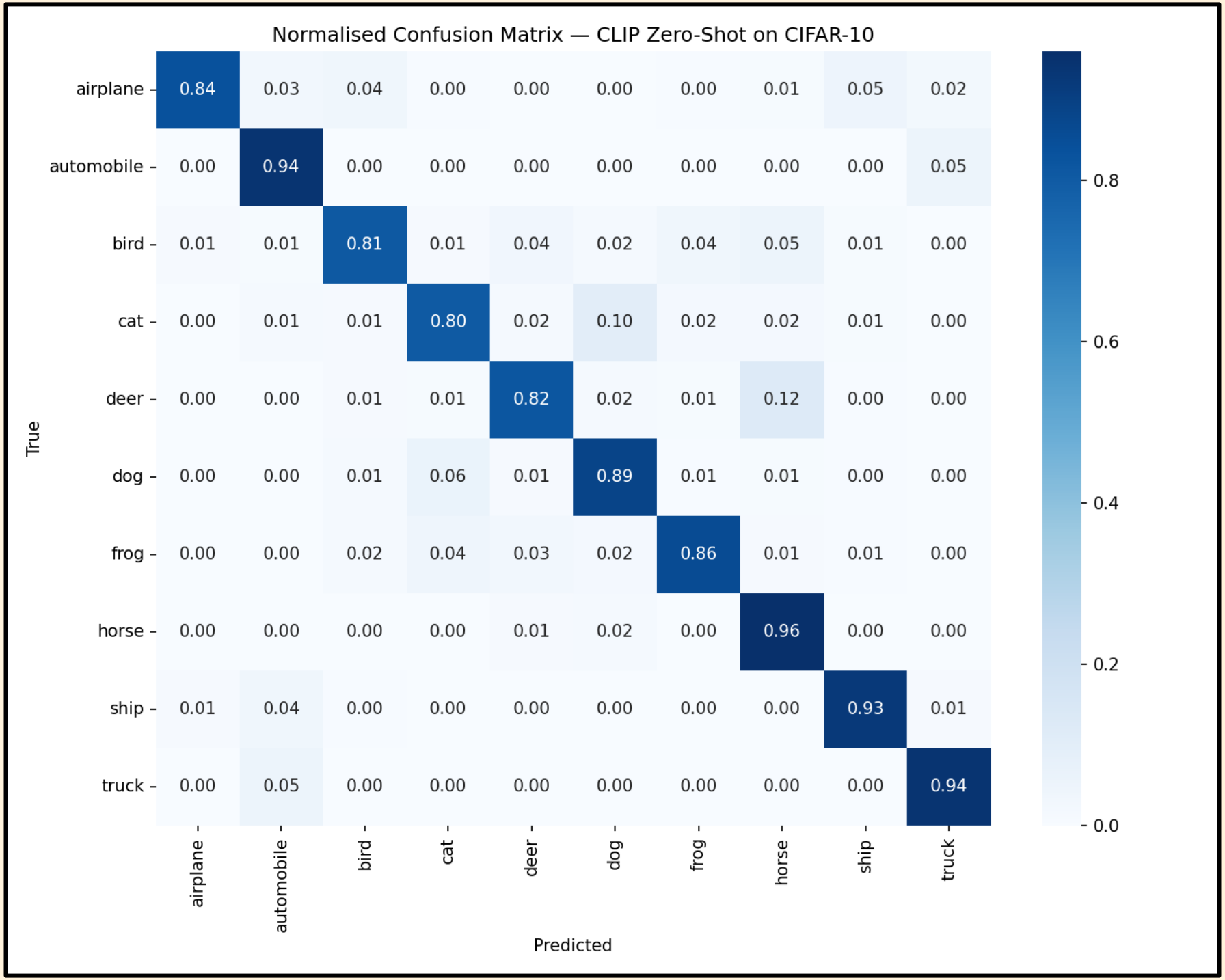


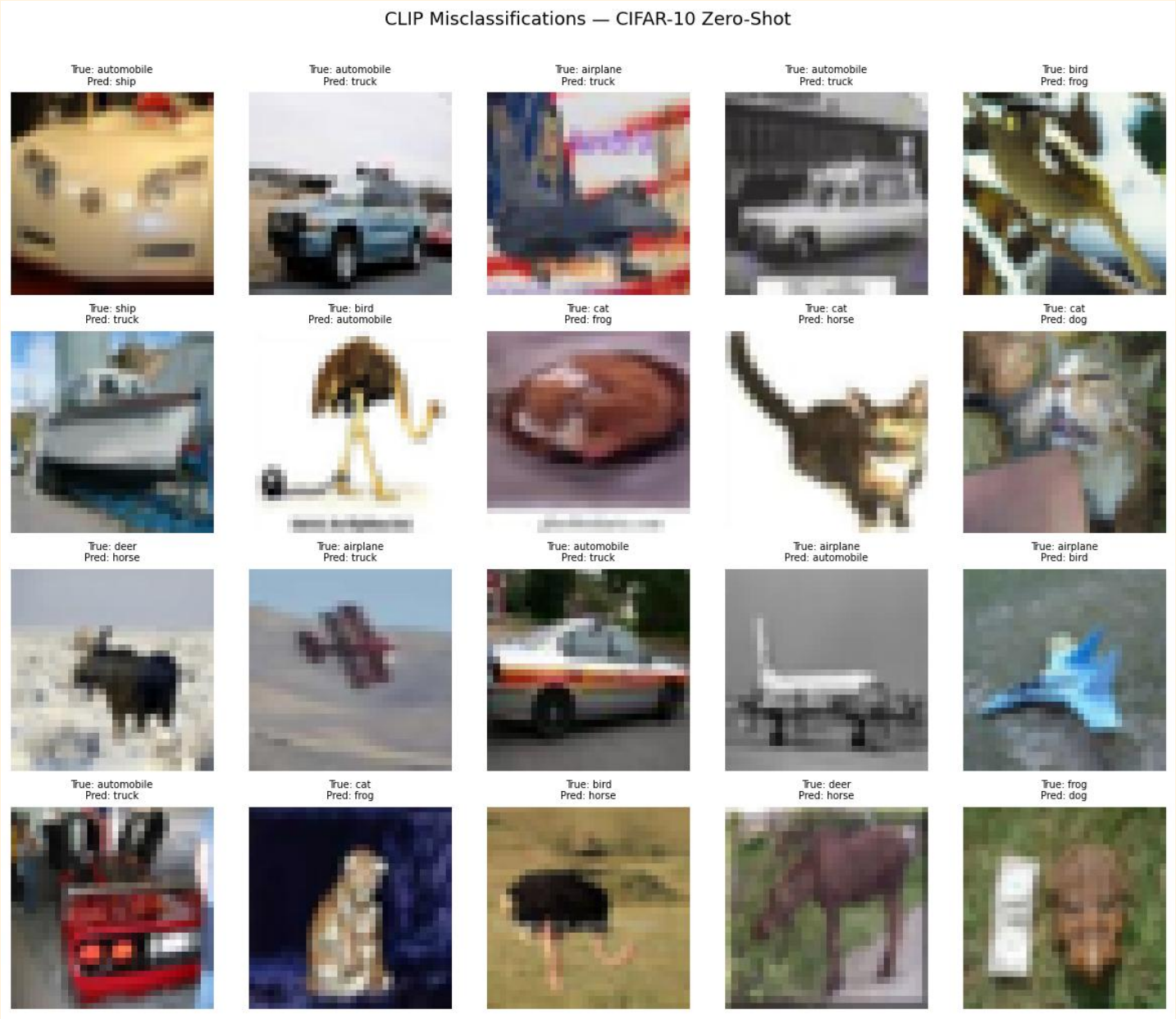


Two Translators: ViT for encoding images and transformer for words.

The Map: Translates the image and the words into points on a shared mathematical embedding space.

The Goal: If the text matches the picture, their points sit close together on the map. If they do not match, their points are pushed far apart.


Normalised Confusion Matrix — CLIP Zero-Shot on CIFAR-10
airplane 0.84 0.03 0.04 0.00 0.00 0.00 0.00 0.01 0.05 0.02
automobile 0.00 0.94 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.05
bird 0.01 0.01 0.81 0.01 0.04 0.02 0.04 0.05 0.01 0.00
cat 0.00 0.01 0.01 0.80 0.02 0.10 0.02 0.02 0.01 0.00
deer 0.00 0.00 0.01 0.01 0.82 0.02 0.01 0.12 0.00 0.00
dog 0.00 0.00 0.01 0.06 0.01 0.89 0.01 0.01 0.00 0.00
frog 0.00 0.00 0.02 0.04 0.03 0.02 0.86 0.01 0.01 0.00
horse 0.00 0.00 0.00 0.00 0.01 0.02 0.00 0.96 0.00 0.00
ship 0.01 0.04 0.00 0.00 0.00 0.00 0.00 0.00 0.93 0.01
truck 0.00 0.05 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.94
True
airplane automobile bird cat deer dog frog horse ship truck
Predicted
0.8
0.6
0.4
0.2
0.0


CLIP Misclassifications — CIFAR-10 Zero-Shot
True: automobile
Pred: ship
True: automobile
Pred: truck
True: airplane
Pred: truck
True: automobile
Pred: truck
True: bird
Pred: frog
True: ship
Pred: truck
True: bird
Pred: automobile
True: cat
Pred: frog
True: cat
Pred: horse
True: cat
Pred: dog
True: deer
Pred: horse
True: airplane
Pred: truck
True: automobile
Pred: truck
True: airplane
Pred: automobile
True: airplane
Pred: bird
True: automobile
Pred: truck
True: cat
Pred: frog
True: bird
Pred: horse
True: deer
Pred: horse
True: frog
Pred: dog

What's next?

- Analysis for Assignment 2
- Research Project

Project 3: Bridging Structure and Function in Glaucoma: A Systematic Study of Deep Learning for Visual Field Prediction from Retinal Imaging

Project 8: Evaluating the Robustness of MLLM Safety Alignment under Image Corruptions



Any Questions?