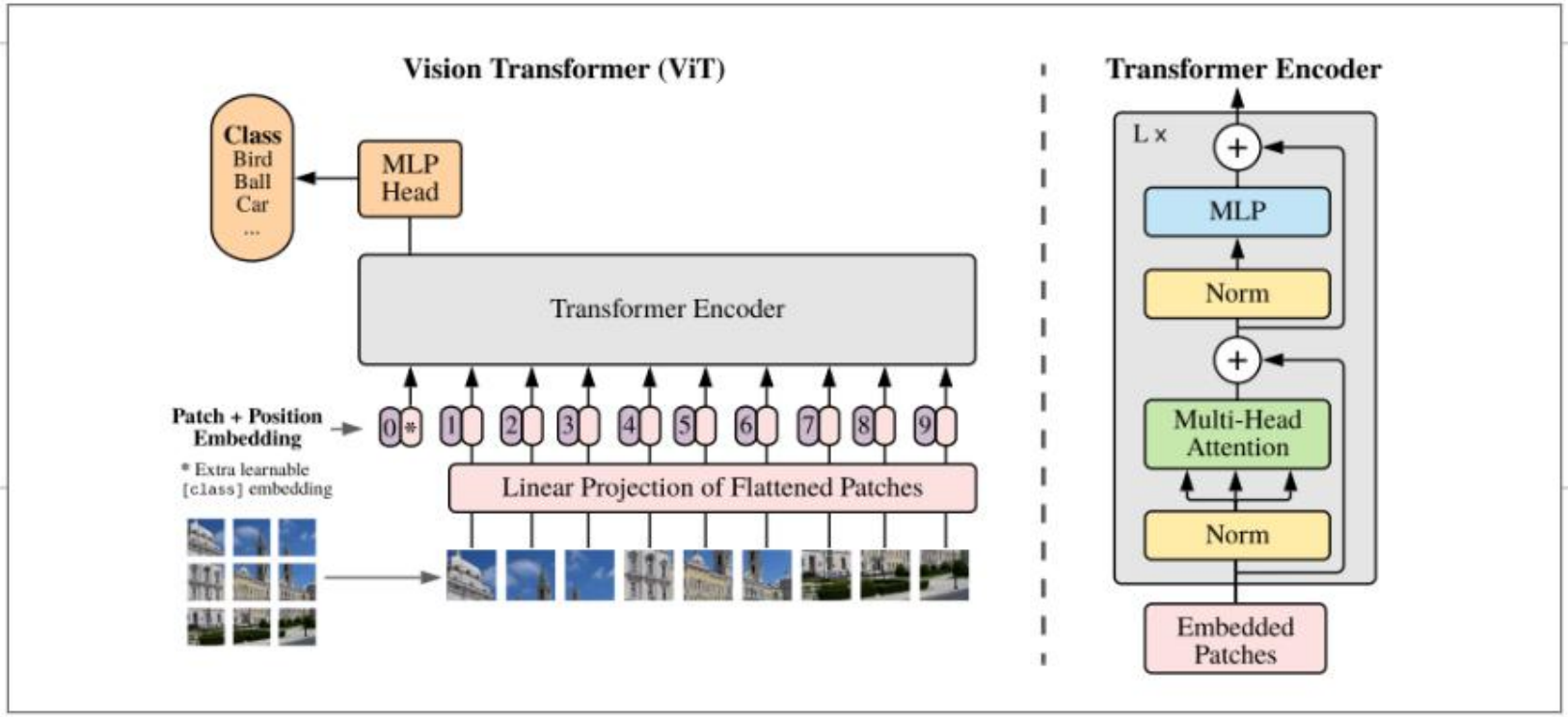
Week 2

David Tang

About Me

- University of Central Florida
- Mathematics Major, Computer Science Minor
- Rising Senior
- Ongoing Honors Thesis on Multi Objective Nash Alignment
  Advised by Professor Yue Wang

Assignment 1 – Vision Transformers

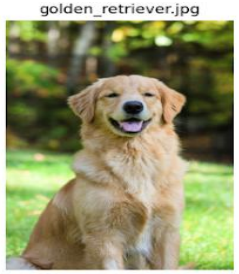


Credit: Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," ICLR 2021.

Assignment 1 – Vision Transformers

#	File	Top-1 (conf.)	Top-2 (conf.)	Top-3 (conf.)
1	red_car.jpg	sports car, sport car 96.46%	racer, race car, racing car 1.58%	car wheel 1.54%
2	rock_landscape.jpg	valley, vale 54.68%	cliff, drop, drop-off 6.16%	jeep, landrover 2.12%
3	mountain_bike.jpg	mountain bike, all-terrain bike, off-roader 65.26%	moped 20.32%	bicycle-built-for-two, tandem bicycle, tandem 4.72%
4	golden_retriever.jpg	golden retriever 96.74%	Labrador retriever 0.78%	tennis ball 0.43%
5	tabby_cat.jpg	tabby, tabby cat 49.58%	tiger cat 35.67%	Egyptian cat 13.63%
6	library_books.jpg	library 70.70%	bookcase 18.42%	bookshop, bookstore, bookstall 9.94%
7	pepperoni_pizza.jpg	pizza, pizza pie 99.34%	dough 0.04%	bakery, bakeshop, bakehouse 0.03%
8	flying_bird.jpg	black stork, Ciconia nigra 31.11%	kite 29.67%	bald eagle, American eagle, Haliaeetus leucocephalus 22.69%
9	airplane.jpg	airliner 98.81%	wing 0.83%	warplane, military plane 0.11%
10	living_room.jpg	studio couch, day bed 80.60%	window shade 4.88%	table lamp 1.35%

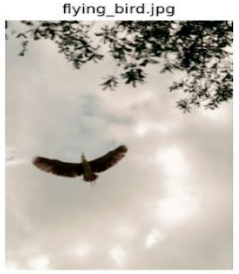
golden_retriever.jpg



tabby_cat.jpg



flying_bird.jpg



red_car.jpg



airplane.jpg



mountain_bike.jpg



pepperoni_pizza.jpg



rock_landscape.jpg



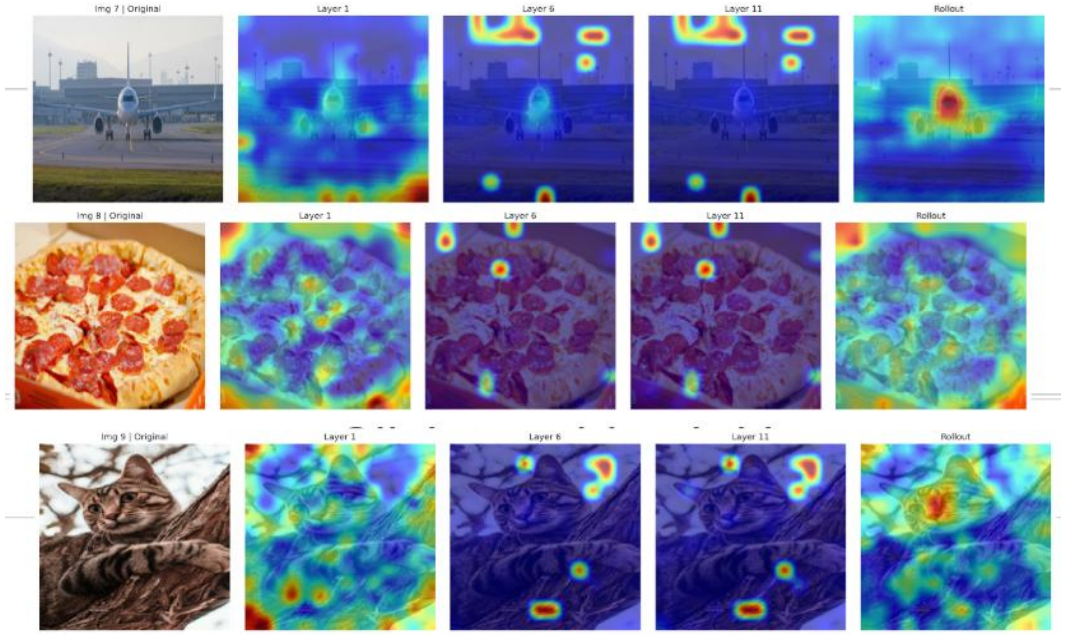
living_room.jpg

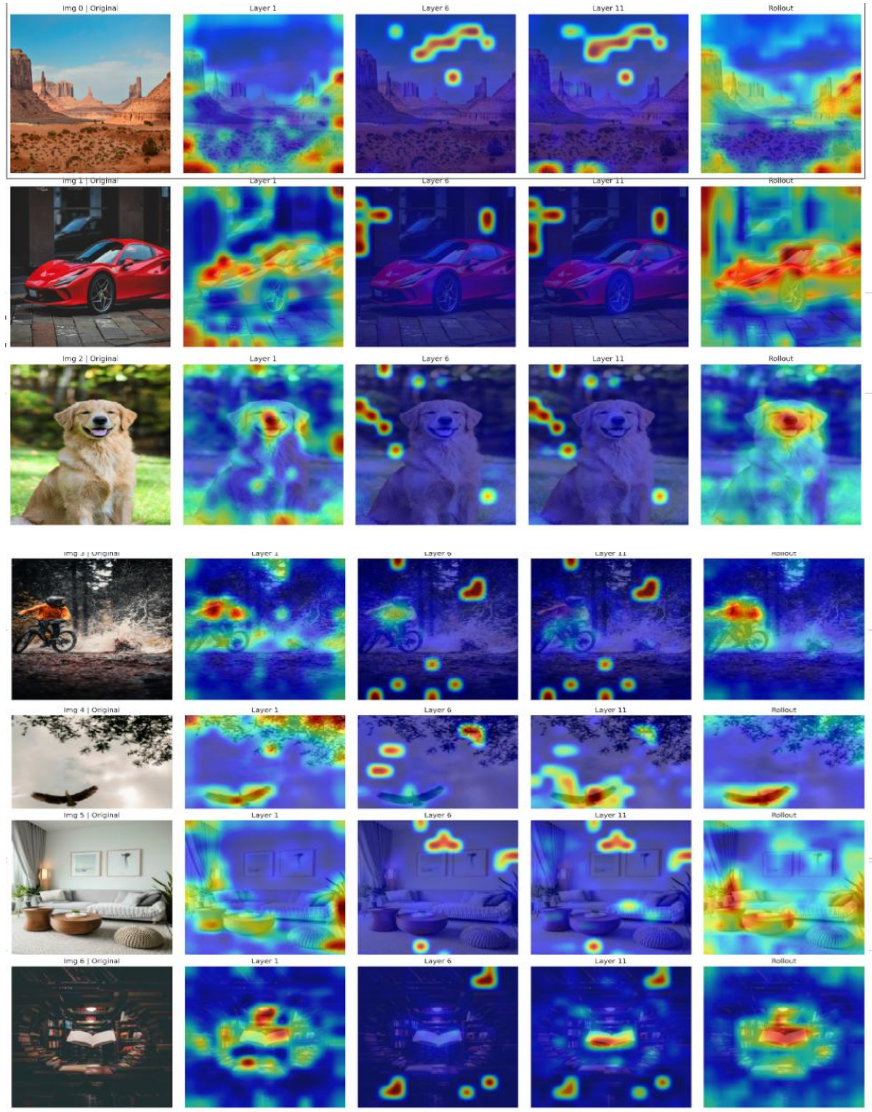


library_books.jpg

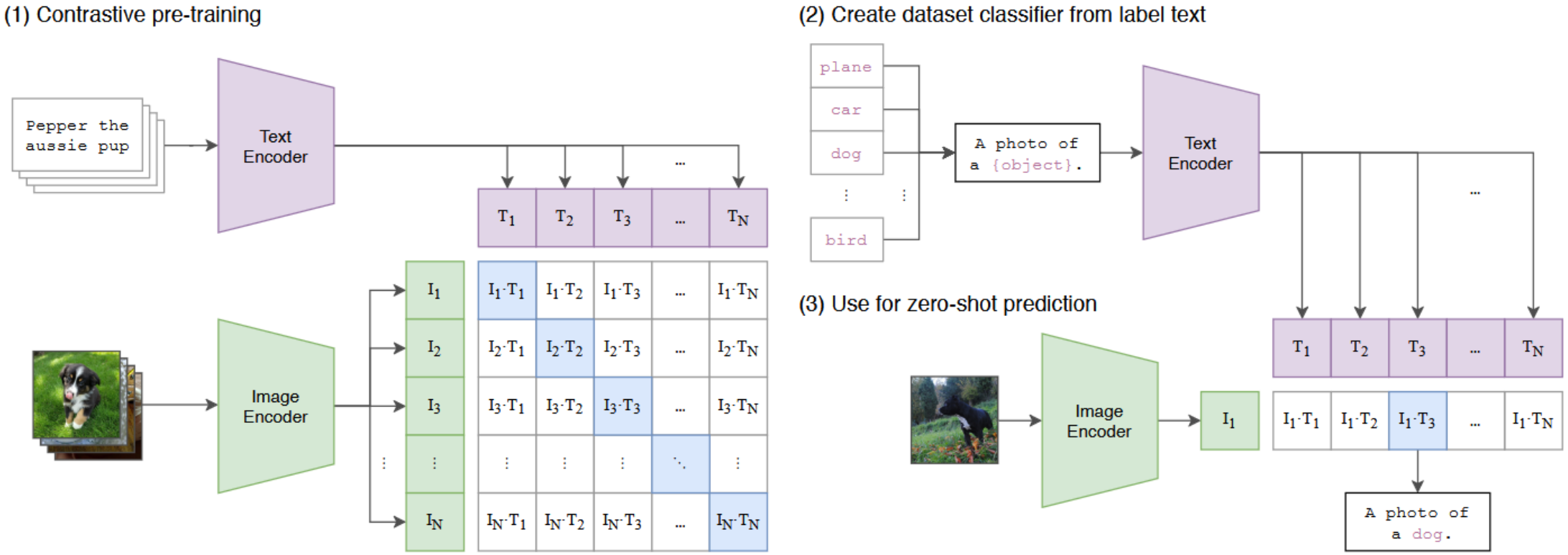


Assignment 1 – Vision Transformers

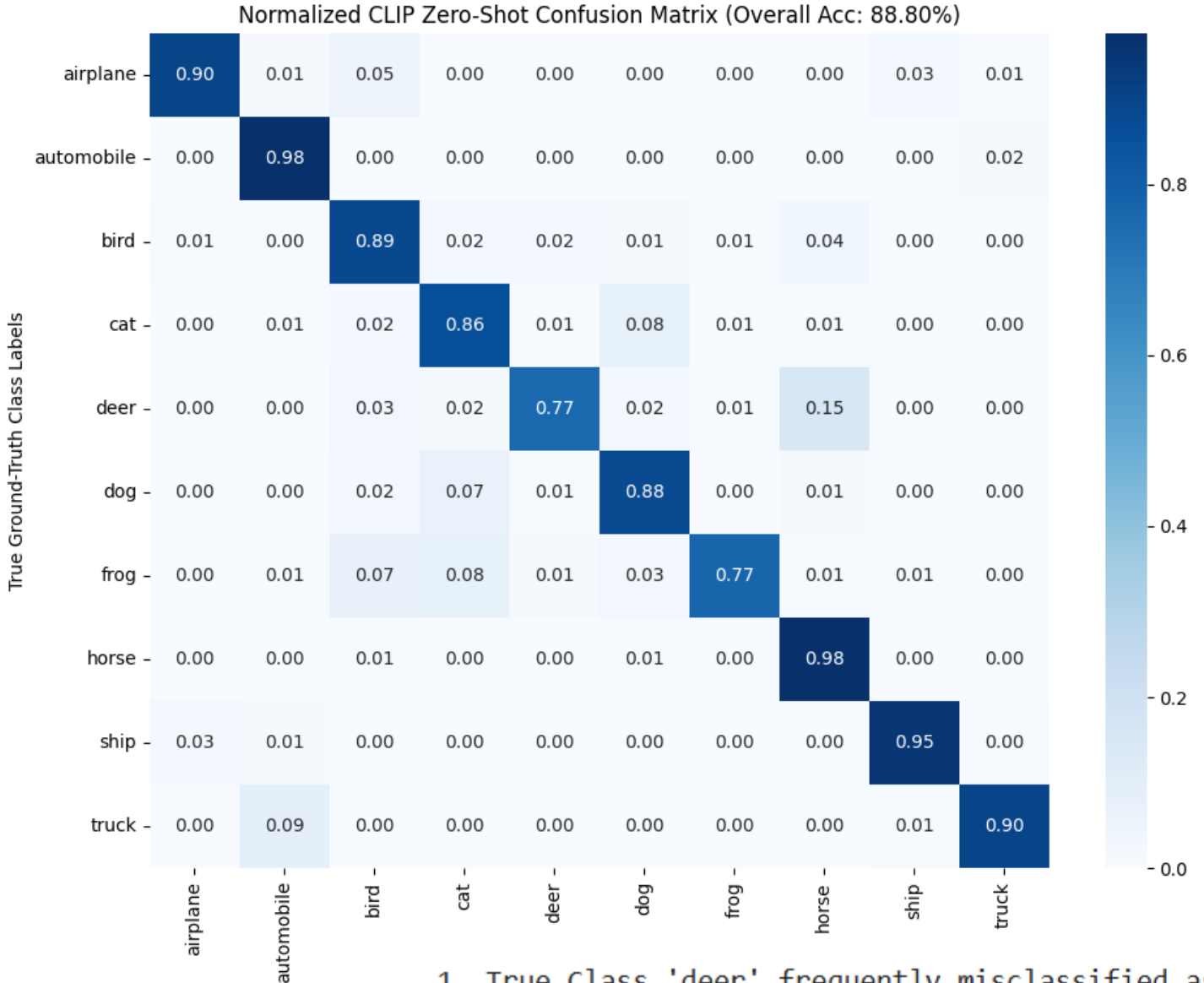




Assignment 2- CLIP



Credit: "Learning Transferable Visual Models From Natural Language Supervision," by A. Radford et al., 2021, arXiv preprint arXiv:2103.00020.



Class	Top-1 Accuracy (%)
airplane	90.1
automobile	97.6
bird	88.8
cat	86.0
deer	76.5
dog	88.4
frog	77.4
horse	97.6
ship	95.5
truck	90.1
OVERALL TOTAL	88.8



1. True Class 'deer' frequently misclassified as Predicted Class 'horse' (Value: 0.1540)
2. True Class 'truck' frequently misclassified as Predicted Class 'automobile' (Value: 0.0890)
3. True Class 'cat' frequently misclassified as Predicted Class 'dog' (Value: 0.0790)

Next Steps

- Finish Up Assignment 2
- Get More Familiar with the cluster
- Refactor and Clean the Codebase