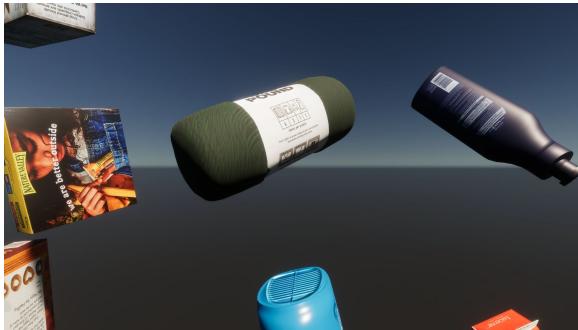
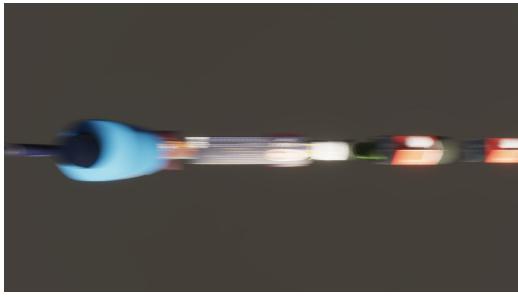
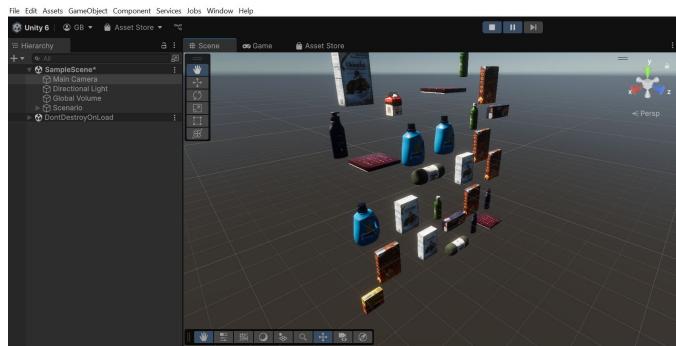


RGB Camera Prefab and Image Capture



*Used Unity Perception Package



Training Computer Vision Model Using RGB Images

Method:

- Used Unity's SynthDet to save and upload 1253 labeled images on Google Drive
- Converted image annotations to YOLO format and split the data in an 8:2 ratio for train and val, respectively
- Ran an initial training on YOLOv8 for 120 epochs with 10 frozen layers, and later fine tuned for 33 more epochs with no frozen layers
- Used Weights and Biases to log results and visualizations

Metrics

Final:

- mAP0.5
 - 0.248
- mAP0.5-0.95
 - 0.191
- Precision
 - 0.287
- Recall
 - 0.248

Epoch GPU_mem box_loss cls_loss dfl_loss Instances Size
33/100 0G 1.15 2.532 1.094 216 416: 100% [██████████] 24/24 [03:50<00:00, 9.61s/it]

Class Images Instances Box(P R mAP50 mAP50-95): 100% [██████████] 3/3 [00:53<00:00, 17.80s/it]

Class	Images	Instances	Box(P)	R	mAP50	mAP50-95: 100%	12/12 [00:52<00:00, 4.34s/it]
all	94	1417	0.277	0.287	0.248	0.191	
book_dorkdiaries_aladdin	15	15	0.17	0.267	0.145	0.123	
candy_minipralines_lindt	25	26	0.555	0.615	0.69	0.503	
candy_raffaeleo_confectioneria	20	22	0.292	0.636	0.482	0.4	
cereal_capt_crunch	15	16	0.376	0.5	0.306	0.276	
cereal_cheerios_honeynut	17	18	0.193	0.111	0.161	0.125	
cereal_corn_flakes	20	21	0.159	0.19	0.0999	0.0722	
cereal_cracklinatbran_kelloggs	17	19	1	0	0.084	0.0625	
cereal_oatmealsquares_quaker	20	23	0.341	0.217	0.173	0.149	
cereal_puffins_barbaras	25	27	0	0	0.0649	0.0523	
cereal_raisin Bran	19	22	0.455	0.182	0.261	0.224	
cereal_rice_krispies	26	29	0.411	0.448	0.435	0.35	
chips_gardensalsa_sunchips	19	20	0.191	0.45	0.308	0.263	
chips_sourcream_lays	17	20	0.368	0.25	0.246	0.177	
cleaning_frequent tide	18	23	0.5	0.739	0.752	0.624	
cleaning_smuggle_henkel	16	17	0.317	0.471	0.348	0.247	
cracker_honeymaid_nabisco	19	22	0.262	0.545	0.269	0.235	
cracker_lightrye_wasa	16	24	0.305	0.329	0.217	0.187	
cracker_triscuit_avocado	22	24	0.178	0.208	0.146	0.129	
cracker_zwieback_brandt	18	21	0.472	0.476	0.434	0.348	
craft_yarn_caron	22	23	0.37	0.522	0.555	0.409	
drink_adrenaline_shock	26	30	0.345	0.467	0.433	0.31	
drink_arabica_beans_kickinghorse	16	17	0.422	0.647	0.554	0.4	
drink_greentea_itoen	17	18	0.419	0.202	0.227	0.136	
drink_orangejuice_minuteman	22	26	0.36	0.231	0.261	0.222	
drink_whippingcream_lucerne	24	27	0.47	0.37	0.42	0.324	
footware_slippers_disney	20	26	0.691	0.654	0.677	0.49	
hygiene_poise_pads	20	21	0.45	0.429	0.409	0.355	
lotion_essentially_nivea	24	24	0.407	0.375	0.316	0.21	
lotion_vanilla_nivea	17	20	0.354	0.35	0.34	0.262	
pasta_lasagna_barilla	23	29	0.14	0.0345	0.0802	0.0479	
pest_antbaits_terro	24	25	0.856	0.04	0.132	0.0584	
porridge_grits_quaker	18	22	0.532	0.136	0.142	0.124	
seasoning_canesugar_candh	12	13	0.215	0.421	0.273	0.145	
snack_biscotti_giotti	23	26	0.249	0.308	0.293	0.249	
snack_breadsticks_nutella	16	17	0.0717	0.0548	0.0529	0.0413	
snack_chips_pringles	26	29	0.333	0.586	0.44	0.314	
snack_coffeecakes_hostess	23	25	0.162	0.12	0.124	0.112	
snack_cookie_famousamos	18	22	0	0	0.0841	0.0704	
snack_cookie_petitecolier	14	15	0.159	0.0667	0.0562	0.0321	
snack_cookie_quadratini	22	26	0.25	0.231	0.165	0.14	
snack_cookie_wafflettes	22	24	0.298	0.333	0.24	0.172	
snack_cookie_walkers	21	25	0	0	0.0734	0.0387	
snack_cookies_foure	20	23	0.315	0.435	0.362	0.304	
snack_granolabar_kashi	25	26	0.243	0.154	0.195	0.151	
snack_granolabar_kind	16	16	0.292	0.375	0.208	0.132	
snack_granolabar_naturevalley	22	24	0.127	0.0417	0.119	0.0993	
snack_granolabar_quaker	20	22	0.223	0.274	0.169	0.147	
snack_salame_hillshire	20	22	0.415	0.0455	0.161	0.115	
soup_chickenenchilada_progresso	20	23	0.352	0.652	0.571	0.48	
soup_tomato_pacific	23	25	0.242	0.12	0.104	0.0874	
storage_ziploc_sandwich	13	13	0.283	0.538	0.429	0.363	
toiletry_tissue_softly	26	27	0.333	0.426	0.343	0.276	
toiletry_toothpaste_colgate	22	24	0.36	0.292	0.239	0.169	
toy_cat_melissa	19	24	0.682	0.625	0.626	0.352	
utensil_candle_decorators	23	29	0.568	0.69	0.729	0.533	
utensil_coffee_filters	20	23	0.576	0.177	0.41	0.31	
utensil_cottonovals_signaturecare	20	23	0.558	0.549	0.505	0.404	
utensil_papertowels_valuecorner	17	17	0.308	0.824	0.498	0.439	
utensil_toiletpaper_scott	20	22	0.435	0.818	0.787	0.694	
utensil_trashbag_valuecorner	24	29	0.307	0.448	0.366	0.204	
vitamin_centrumsilver_adults	14	17	0.108	0.235	0.0619	0.0504	
vitamin_centrumsilver_men	18	21	0.0704	0.0952	0.0567	0.0488	
vitamin_centrumsilver_woman	26	28	0.198	0.214	0.143	0.114	

Confusion Matrix

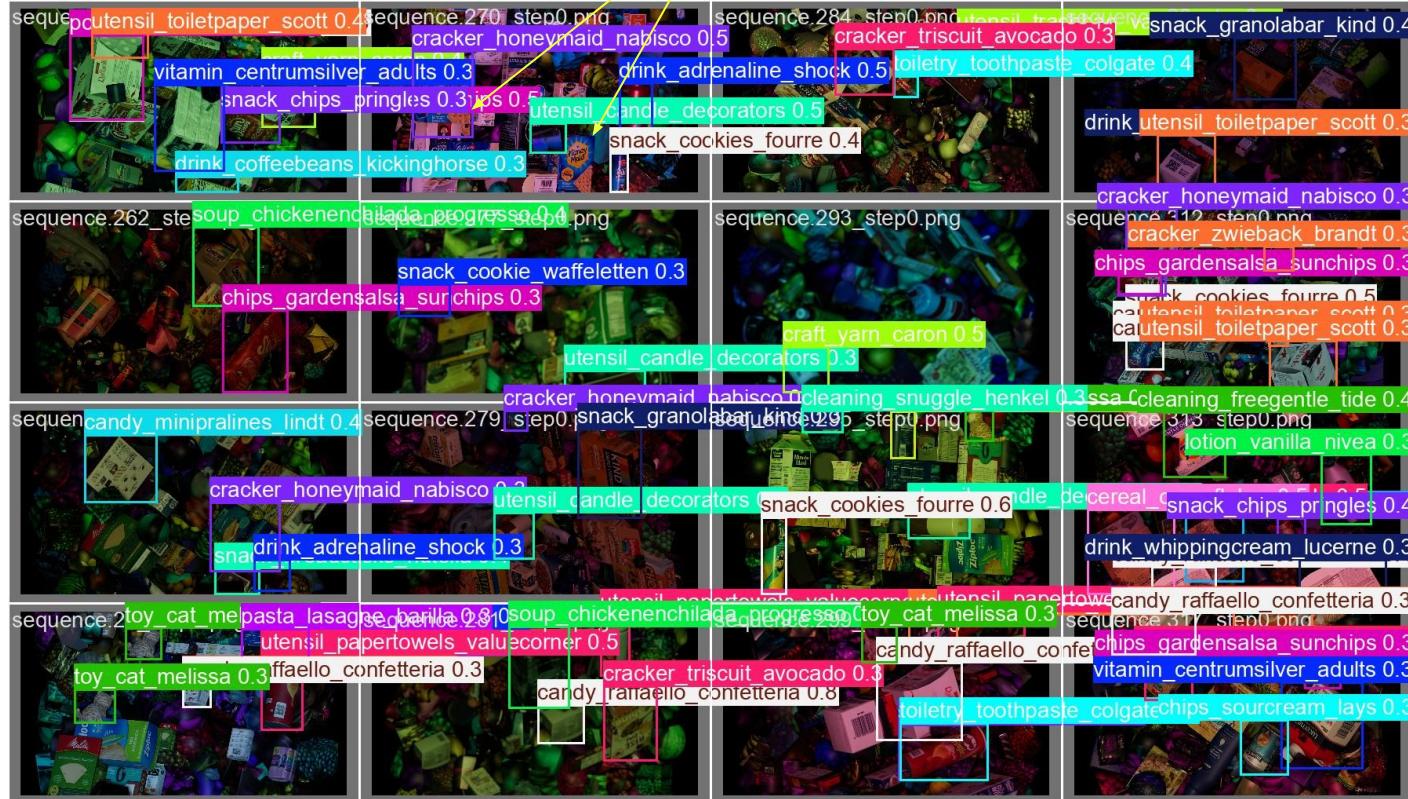
- 94 validation images
- 1417 instances
- 63 classes



Detection Images

Model is learning, but still needs improvement (missed an object from a previously detected class)

Step 1



Detection Images



Step 2

Detection Images

Model mistook a food item for a toiletry product



Step 3

Detection Images



Step 4

Model is able to detect objects from convoluted images

Detection Images

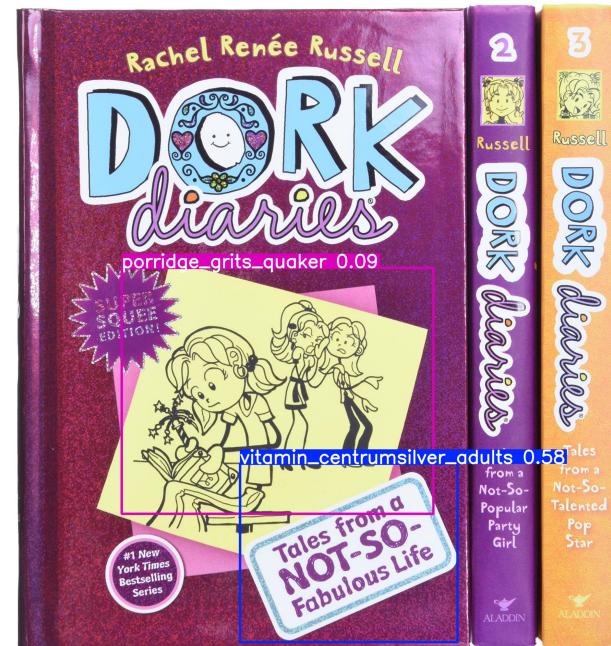
Step 5

*That being said, training is very time intensive.

Will consult for next steps (regarding additional fine tuning for underperforming classes or using model for a new assignment)

Real Image Detection Using SynDet Trained Model

- No correct detections across all 63 classes before fine-tuning with real images
- *Considering that model was trained on synthetic data, this is not surprising
 - Model needs real images to account for different factors such as lighting and background



Methods - Finetune w/ Real Images

- Saved and uploaded 604 real images on Google Drive to finetune object detection model classes
- Converted image annotations to YOLO format and split the data in an 8:2 ratio for train and val, respectively
 - Used full image bounding boxes (each entire image corresponds to a single class)
- Ran 30 epochs from synthetic data-only pretrained model (best.pt) with 1 frozen layer
- Used Weights and Biases to log results and visualizations for each class

Metrics - Finetune w Real Images

- Underperforming classes show 0 recall
- Well performing classes have high mAP50 though most only appear in 1 instance
- mAP 50
 - 0.2522519578547975
- mAP 50-0.95
 - 0.2490688194471003
- Precision
 - 0.28928928383898883
- Recall
 - 0.2637652351871677

Class	Images	Instances	Box/P	R	mAP50	mAP50-95: 100%		
[00:14<00:00, 1.86s/it]								8/8
all	120	120	0.289	0.264	0.252	0.249		
book_dorkdiaries_aladdin	4	4	0.23	0.383	0.37	0.37		
candle_decorators	4	4	0.163	0.5	0.354	0.354		
cereal_capn_crunch	1	1	0	0	0.0262	0.0262		
cereal_cheerios_honeynut	2	2	0.107	1	0.39	0.39		
cereal_com_flakes	1	1	0	0	0.0765	0.0765		
cereal_cracklinotabran_kelloggs	2	2	1	0	0.034	0.034		
cereal_oatmealsquares_quaker	1	1	0	0	0.0355	0.0355		
cereal_puffins_barbaras	1	1	0.11	1	0.995	0.995		
cereal raisin bran	3	3	0.216	0.667	0.515	0.515		
chips_gardensalsa_sunchips	3	3	0.272	0.272	0.357	0.357		
chips_sourcream_lays	1	1	0	0	0.0131	0.0131		
cleaning_freetgentle_tide	2	2	0	0	0.194	0.0751		
coffee_filters	4	4	0.316	0.25	0.396	0.396		
cottonovals_signaturecare	2	2	0.0858	0.5	0.0939	0.0939		
cracker_honeymaid_nabisco	2	2	0	0	0.0243	0.0243		
cracker_lightryte_wasa	1	1	0	0	0.0231	0.0231		
cracker_triscuit_avocado	1	1	1	0	0.0711	0.0711		
cracker_zwieback_brandt	3	3	0.398	1	0.497	0.497		
craft_yarn_caron	2	2	0	0	0.00894	0.00894		
drink_adrenaline_shock	3	3	1	0	0.696	0.696		
drink_coffeebeans_kickinghorse	7	7	0	0	0	0.207		0.207
drink_greentea_itoen	2	2	0.145	0.5	0.249	0.249		
drink_whippingcream_lucerne	1	1	0	0	0	0.0332		0.0332
footware_slippers_disney	1	1	1	0	0	0		
hygiene_poise_pads	2	2	0	0	0.133	0.133		
lotion_essentialy_nivea	3	3	0.442	0.667	0.398	0.398		
lotion_vanilla_nivea	1	1	0.182	1	0.995	0.995		
papertowels_valuecorner	2	2	1	0	0.152	0.152		
pasta_lasagne_barilla	1	1	0.173	0.867	0.199	0.199		
pest_antbaits_terro	4	4	0.279	0.75	0.79	0.79		
porridge_grits_quaker	2	2	0	0	0.0714	0.0714		
seasoning_canesugar_candh	1	1	0	0	0.0149	0.0149		
snack_biscotti_ghiotto	3	3	1	0	0.137	0.137		
snack_breadsticks_nutella	2	2	0.19	0.5	0.595	0.595		
snack_chips_pringles	2	2	0	0	0.0226	0.0226		
snack_coffeecakes_hostess	3	3	0.274	0.519	0.453	0.453		
snack_cookie_famousamos	2	2	0	0	0.0524	0.0524		
snack_cookie_petitecolier	1	1	0	0	0.0415	0.0415		
snack_cookie_quadratini	3	3	0.085	0.333	0.393	0.393		
snack_cookie_waffelletten	3	3	0.178	0.333	0.102	0.102		
snack_cookie_walkers	1	1	0	0	0.00975	0.00975		
snack_cookies_fourre	4	4	0	0	0.0787	0.0787		
snack_granolabar_kashi	3	3	1	0	0.0563	0.0563		
snack_granolabar_kind	1	1	1	0	0.124	0.124		
snack_granolabar_naturevalley	1	1	0	0	0.0269	0.0269		
snack_granolabar_quaker	1	1	0	0	0.0221	0.0221		
snack_salamie_hillshire	3	3	0.326	1	0.995	0.995		
soup_chickenenchilada_progresso	3	3	0.2	0.133	0.237	0.237		
storage_ziploc_sandwich	1	1	0	0	0.0112	0.0112		
tissue_softly	3	3	1	0	0.0889	0.0889		
toiletpaper_scott	2	2	1	0	0.308	0.308		
toothpaste_colgate	1	1	0.193	1	0.995	0.995		
trashbag_valuecorner	3	3	1	0	0.15	0.15		
vitamin_centrumsilver_men	1	1	0.188	1	0.332	0.332		
vitamin_centrumsilver_woman	3	3	0.159	0.333	0.228	0.172		

Confusion Matrix

- Poor accuracy with several misclassifications across classes
 - Background class is frequently predicted



0.25 Confidence Detections

- Low detection rates
- *No detections for high performing classes (ex. puffin cereal)



predictions_snack_salame_hillshire



0.1 Confidence Detections

- Increased accuracy and correct detections
- Vulnerable to false positives



Limitations

- Smaller sample size
 - ~10 real images per class
 - ~2 images in validation set per class
- Fewer epochs
 - 30 epochs compared to ~150
- Full images were used as bounding boxes

.stl training

- Classes
 - Coffee filters and *Detergent Bottles
- Limitations
 - Money
 - Limited to thingiverse and printables for free content and very few classes had viable .stl files
 - Detail
 - .stls were not as detailed or true to the synthetic data and (byproduct of above problem)

.stl Training Methods

-

.stl Classes

Coffee Filters -

- <https://www.printables.com/model/1238160-coffee-filter-basket/files>
- <https://www.printables.com/model/984996-kalita-wave-155-dripper-coffee-filter-crimp/files>
- <https://www.thingiverse.com/thing:4419392/files>
- <https://www.thingiverse.com/thing:5344645/files>

Detergent Bottle*

-

Next Steps

- Train underperforming classes with more real images
 - Tailor bounding box labels for each image
-