

به نام خدا



دانشگاه تهران دانشکده مهندسی برق و کامپیوتر شبکه های عصبی و یادگیری عمیق

تمرین امتیازی

ژبیوار صورتی حسنزاده	نام و نام خانوادگی		
810196502	شماره دانشجویی		
1400/03/6	تاریخ ارسال گزارش		

فهرست

Object Detection With YOLOv5
Semantic Segmentation

4

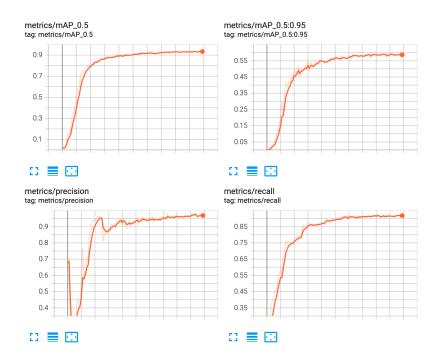
Object Detection With YOLOv5

۱. در مورد شبکه yolov4 میتوان گفت این شبکه از cspdarknet53 به عنوان backbone استفاده میکند که به نوعی feature map های مربوطه را PANet path-aggregation به عنوان الله آخر که extract به استفاده کرده و به عنوان لایه آخر که head باشد از head مانند yolov3 استفاده میکند که میتوان گفت پیشرفتی در این قسمت از آن spatial pyramid بهبود عملکرد backbone و به طور به خصوص pooling های آن از image به طور pooling استفاده شده است. بهبود عملکرد به این شکل انجام میشود که با استفاده از spp ها، یک بار image به طور کامل pooling استفاده از pooling های متفاوت عمل تبدیل آن به dimension تعیین شده در استفاده از fully کامل connected layer های میشود که برای همان region های متفاوت میباشد. همینطور میتوان گفت با استفاده از bof ها و bos ها و bos ها و backbone و استفاده از poling های مستند مانند augmentation و استفاده از will استفاده از mish activation function انجام میشود که برای همان poling های متفاوت میباشد. همینطور بیشرفتی که در yolov5 انجام شد به شکل استفاده از learning bonding box anchors و مینطور پیشرفت بسیار قابل ملاحظه ای در زمینه زمان processing ایجاد کرد که به میزان yersion های و بود که در مقایسه با processing دیده شده در yolov4 بود.

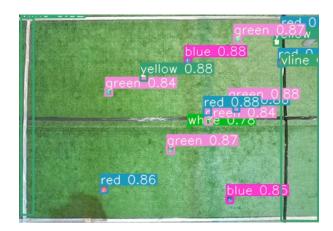
٠٢

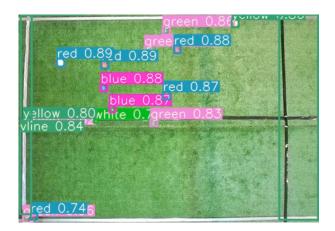
Epoch	gpu_mem	box	obj	cls	total	. targets	img_size	
99/99	1.81G	0.04915	0.04042	0.008595	0.09816	270	416:	100% 99/99
	Class	Images	Targets		Р	R	mAP@.5	mAP@.5:.95
	all	58	1.00e	+03	0.966	0.92	0.935	0.583
	blue	58		115	1	0.97	0.995	0.607
	green	58		290	0.993	0.953	0.989	0.579
	red	58		290	0.993	0.993	0.996	0.657
	vline	58		136	0.967	0.978	0.978	0.851
	white	58		58	0.841	0.638	0.655	0.189
	yellow	58		116	1	0.99	0.995	0.617

اطلاعات مربوط به آموزش برای epoch آخر



نمودارهای metric های مربوط به آموزش مدل نسبت به metric و recal





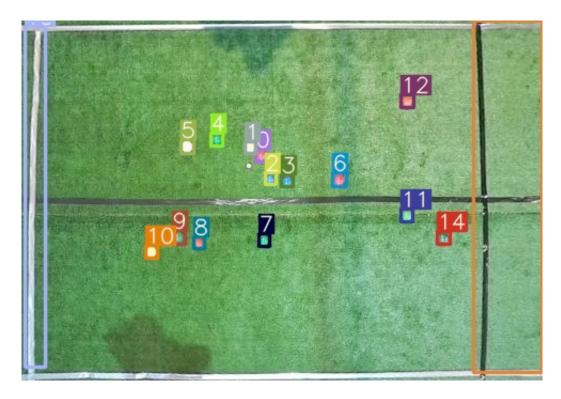
توپهای شناسایی شده به همراه confidence مربوطه

با بررسی چند تصویر مربوط به دادههای تست مشاهده میشود با وجود اینکه مدل با accuracy کامل و %100 عمل نمیکند ولی به طور کلی عملکرد خوبی روی دادههای تست دارد. همینطور برای آنکه کیفیت تصویرهای بدست آمده را بهتر کنیم در قسمت line thickness در قایل detect.py این پارامتر را کمتر از حالت اصلی آن که 3 بود قرار دادیم تا خوانایی تصویر بیشتر شود.

.۴



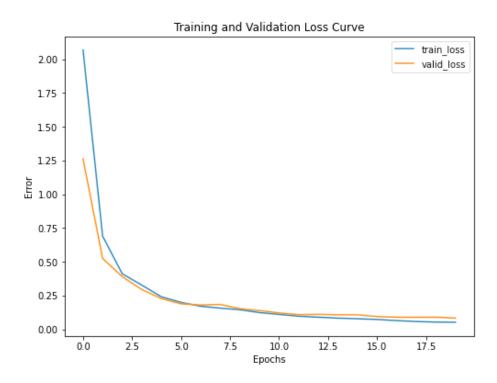
توپهای شناسایی شده با index مربوط به نزدیکی آنها نسبت به توپ سفید



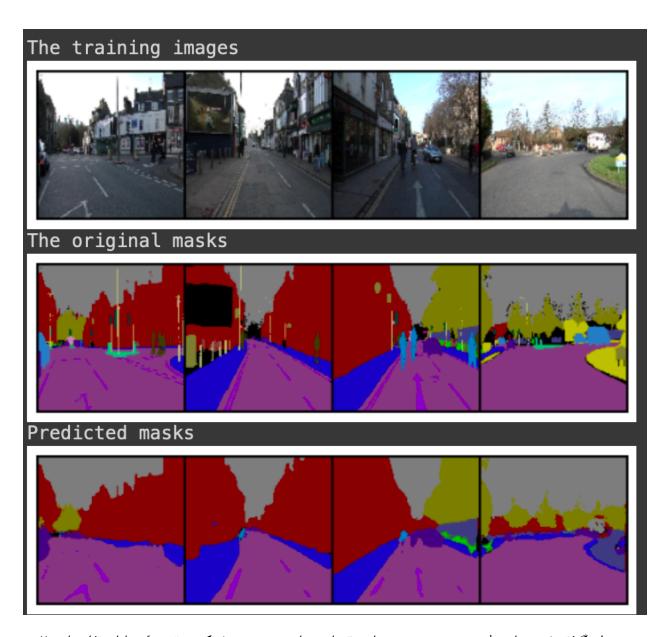
توپهای شناسایی شده با index مربوط به نزدیکی آنها نسبت به توپ سفید

Semantic Segmentation

۱) در شکل زیر میتوانید نمودار خطا مربوط به دادههای اَموزش و validation را مشاهده نمایید.



Loss مربوط به دادههای آموزش و Loss



در نظر گرفته شده برای epoch خروجی مربوط به تعدادی داده ورودی به شبکه و نتیجه آن با استفاده از ۲۰ training

۳) با توجه به تعداد run های انجام شده و دقت به نمودار loss مربوط به validation میتوان مشاهده کرد که نمودار مربوطه با انجام ۴۰ epoch بنیز به کم شدن loss ادامه می دهد که البته با توجه به سنگین بودن شبکه تعداد بیشتر از این مقدار در نظر گرفته نشد و مقدار ۴۰ را برای این منظور در نظر میگیریم. همینطور برای استفاده از دادههای تست و آموزش نیز از دادههای خود که train و Kaggle استفاده شده از Kaggle استفاده شده از به جداسازی مجدد نبود. اما به طور کلی برای جداسازی دادههای مربوط به classification بهتر است به نوعی دادهها تقسیمبندی شوند که category های موجود در هر دو قسمت داده به طور balanced وجود داشته باشند.

در این شبکه از loss ای به نام Dice Loss layer استفاده می شود. هدف استفاده از این تابع این است که با توجه به اینکه در تصاویر پزشکی و انجام segmentation در آنها اکثرا این اتفاق می افتد که تصویر فانجام background غالب است، در نتیجه تصویر مربوط به target یا همان تصویری که از background متفاوت است، اغلب شناسایی نشده و یا مقدار بسیار محدودی از آن شناسایی می شود. در پاسخ به این مشکل در VNet از loss function ای به شکل زیر استفاده شده است.

$$D = \frac{2\sum_{i}^{N} p_{i}g_{i}}{\sum_{i}^{N} p_{i}^{2} + \sum_{i}^{N} g_{i}^{2}}$$

 $p_i \in P$: predicted binary segmentation volume

 $g_i \in G$: ground truth labels

نتیجه گرفته شده از استفاده از این تابع loss به این شکل میباشد که دیگر نیازی به در نظر گرفتن وزنهای balance کننده میان قسمتهای متفاوت عکس نمیباشد که در حالت عادی با در نظر گرفتن وزنهایی در loss weight ها در نظر گرفته میشود.