



تمرین سری 7 درس مبانی بینایی کامپیوتر

نام مدرس: دکتر محمدرضا محمدی
دستیار آموزشی مرتبط: پارسا عیسی زاده –
شایان موسوی نیا – مهدی خورشیا
مهلت تحویل: ۱۴۰۲ / تیر / ۱۶

1 - در این سؤال قرار است برای مجموعه داده‌ای که در اختیار دارید مدل yolov7 برای تشخیص پنل‌های خورشیدی train کنید. اما این مجموعه داده دارای برچسب‌های مناسب برای آموزش مدل yolov7 نیست و باید برای آن ساخته شود. (60 نمره)

- درون نوت‌بوکی که به پیوست داده شده است، ابتدا از mask‌های هر تصویر استفاده کنید تا کانتورهای پنل‌های خورشیدی را استخراج کنید.
- سپس مستطیل محاط به هر کانتور را به دست آورید.
- از اطلاعاتی که مستطیل محاط شده به دست می‌دهد استفاده کنید تا بتوانید برچسب‌های مناسب هر تصویر برای مدل yolov7 را تولید و ذخیره کنید. برای این کار در نظر داشته باشید مدل yolo انتظار دارد در کنار هر فایل تصویر آموزشی، یک فایل txt. به همان اسم به عنوان برچسب ببیند که درون آن به‌ازای هر شی درون تصویر یک سطر با اطلاعات زیر وجود داشته باشد:
- Class_id x_center_normalized y_center_normalized w_rect_normalized h_rect_normalized
این اطلاعات با یک space از هم جدا شده‌اند.
- کلاس‌ها از 0 شروع می‌شود. فرض کنید اگر در تصاویرتان 5 کلاس شی دارید، آنگاه idهای شما از 0 تا 4 خواهند بود.
- نحوه normalize کردن مختصات x و y به این صورت است که کافی است مختصات را به تناسب به عرض و طول تصویر تقسیم کنید.
- سپس کد منبع مدل yolov7 را از آدرسی که داده شده است clone کرده و بر اساس گام‌هایی که درون readme آن آورده شده است، مدل را بر روی داده‌های خودآموزش دهید. لطفاً حین آموزش مدل، به تنظیم درست hyperparameterها دقت کنید.

لینک دریافت تصاویر:

https://drive.google.com/file/d/1KWnX3eMPJrzhsegi0LmyGUbUV5pqKw_R/view

2 - همان‌طور که می‌دانید الگوریتم SORT از سه مرحله اصلی تشکیل شده است:

1. پیدا کردن objectها در frame اولی توسط شبکه Faster R-CNN. الگوریتم Faster R-CNN، آبجکت‌ها را در frame اولیه به تعدادی گره، embed می‌کند.
2. فیلتر Kalman حرکت بعدی object را توسط اعمال ریاضی (بدون استفاده از شبکه deep) پیش‌بینی می‌کند.
3. متناظر کردن objectهای جدید به قبلی‌ها (الگوریتم Hungarian)

برای توضیح ریاضی فیلتر Kalman می‌توانید از ویدئوی لینک زیر کمک بگیرید:

<https://www.youtube.com/watch?v=IFeCIbljreY>



تمرین سری 7
درس مبانی بینایی کامپیوتر

نام مدرس: دکتر محمدرضا محمدی
دستیار آموزشی مرتبط: پارسا عیسی زاده –
شایان موسوی نیا – مهدی خورشیا
مهلت تحویل: ۱۴۰۲ / تیر / ۱۶

همچنین از source code های زیر می توانید ایده بگیرید:

الگوریتم SORT :

<https://github.com/abewley/sort>

الگوریتم فیلتر Kalman :

https://github.com/r1abbe/filterpy/blob/master/filterpy/kalman/kalman_filter.py

حال به سؤالات زیر پاسخ دهید (20 نمره):

- اعمال ریاضی فیلتر Kalman را برای پیش بینی 1 پارامتر توضیح دهید.
- اعمال ریاضی Kalman را به شکل ماتریسی برای پیش بینی حرکت labelها (با 4 ویژگی) در شبکه ای که عکس ها را به 7 گره embed می کند توضیح دهید.
- تفاوت الگوریتم Deep SORT با الگوریتم SORT (همان طور که از اسمش معلوم است) این است که از یک شبکه Deep استفاده می کند . به چه منظوری از شبکه Deep استفاده می شود ؟

3 – الف) یکی از مهم ترین بخش های ساخت یک مدل هوش مصنوعی، فراهم کردن مجموعه داده های آموزشی است. در این سؤال، هدف ما برچسب زنی یک سری فریم از یک قفس است که در آن تعدادی پرندۀ قرار دارند که با آن بتوانیم داده های آموزش مورد نیاز خود را بسازیم. برای برچسب زنی، در این سؤال قصد داریم از نرم افزار cvat (<https://www.cvat.ai/>) استفاده کنیم. پس از اتمام برچسب زنی، خروجی برچسب ها را به فرمت yolov5 از cvat استخراج می کنیم و با استفاده از آن، مدل yolov5s را آموزش می دهیم. در این آموزش، batch size برابر با 2 و تعداد epoch ها برابر با 50 خواهد بود. نتایج آموزش و برچسب های زده شده را در پاسخ خود ارسال کنید. (فایل های مورد نیاز این سؤال در لینک زیر موجود است) (40 نمره)

<https://drive.google.com/drive/folders/1m004UAo-h7A5Vt5eTqcrvEYANh9iQ0G4?usp=sharing>

3 – ب) در بخش بالا دیدیم که چگونه با داشتن مجموع داده مناسب، می تواند یک مدل هوش را برای object detection آموزش داد. آیا با آموزش مدل بر روی یک مجموعه داده که پرندگان آن در قفس نیستند نیز می توان مدلی خوبی ساخت؟ در صورت موافق بودن یا مخالف بودن، دلیل خود را بیان کنید. (10 نمره)

3 – ج) در خیلی از حوزه ها، فراهم کردن مجموعه داده برای موارد خاص، سخت خواهد بود و همچنین برچسب زنی بر روی یک مجموعه داده خام، بسیار زمان بر است. راهکاری ارائه دهید که بتوان از آن برای detection پرندگان در قفس، با استفاده از مجموعه داده های آماده (مانند COCO) استفاده کرد که با آن بتوان مدلی را فراهم کرد که پرندگان را در قفس نیز به خوبی بشناسد. (10 نمره)



تمرین سری 7
درس مبانی بینایی کامپیوتر

نام مدرس: دکتر محمدرضا محمدی
دستیار آموزشی مرتبط: پارسا عیسی زاده –
شایان موسوی نیا – مهدی خورشیا
مهلت تحویل: ۱۴۰۲ / تیر / ۱۶

4 - به سؤالات زیر در مورد siamFC پاسخ دهید: (20 نمره)

- مزایا و محدودیت‌های معماری شبکه siamFC در وظایف بینایی کامپیوتر چیست؟
- شبکه سیامی (SiamFC) در ردیابی اشیاء بصری چگونه عمل می‌کند و اجزای اصلی آن چیست؟
- چالش‌های مرتبط با ردیابی شی چیست و SiamFC چگونه آنها را برطرف می‌کند؟
- چگونه مفهوم معماری سیامی فراتر از ردیابی شی گسترش می‌یابد و چه کارهای بینایی کامپیوتری دیگری می‌تواند از آن بهره‌مند شود؟ (به طور مثال، One-Shot Learning یکی از حوزه‌های مورد استفاده این شبکه است).

لطفاً سند قوانین انجام و تحویل تمرین‌های درس را مطالعه و موارد خواسته شده را رعایت کنید.

موفق باشید.