

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



شبکه های عصبی مصنوعی و یادگیری عمیق

تمرین ویژه شماره یک

اردیبهشت ماه 1400

فهرست سوالات

3..... Object Detection with YOLOv5 – 1 سوال

4..... Semantic Segmentation ۲ سوال

سوال 1 – Object Detection with YOLOv5

در این سوال می‌خواهیم استفاده از YOLO، به طور خاص ورژن ۵، را برای حل یک مسئله واقعی ببینیم. برای اینکار از نوت‌بوک منتشرشده توسط ناشر YOLOv5 بر روی گوگل کولب استفاده خواهیم کرد. می‌توانید نوت‌بوک را در [این لینک](#) مشاهده کنید. در کنار فایل صورت پروژه، یک فایل مجموعه داده‌ی لیبل شده نیز در اختیار شما قرار می‌گیرد که برای آموزش و تست مدل باید از آن استفاده کنید. مجموعه داده شامل تصاویر یک بازی توپی است با نام bocce ball که در کلاس‌ها شامل توپ‌های سفید، قرمز، سبز، آبی و زرد به همراه یک کلاس برای خط‌های عمودی زمین خواهد بود.

۱ - ورژن‌های ۱ تا ۳ خانواده مدل‌های YOLO را مرور کردیم. در مورد ورژن ۴ و ۵ از خانواده مدل‌های YOLO تحقیق کنید و پیشرفت‌های آن‌ها را مختصر توضیح دهید.

۲ - با استفاده از نوت‌بوک YOLOv5 و داده‌های آماده‌شده، یک مدل از YOLOv5 را آموزش دهید. قدم به قدم با نوت‌بوک پیش‌بروید و هر قدم را توضیح دهید. دقت کنید با توجه به اینکه داده‌ها آماده هستند شما نیازی به دالود آنها ندارید و تنها باید آن‌ها را extract کنید. تعداد کلاس‌ها در مدل را با توجه به داده‌ها تنظیم کنید. در انتها نتایج آموزش در epoch آخر و نمودارهای $mAP_{0.5}$ Precision و Recall را در گزارش خود بیاورید.

۳ - تصاویر موجود در پوشه‌ی تست را به مدل بدهید و نتایج به دست آمده را بررسی کنید. دو نمونه از خروجی‌ها را در گزارش بیاورید. تلاش کنید با اعمال تغییرات در کد فایل `detect.py` تصاویر به دست آمده را خواناتر کنید.

۴ - در این قسمت باید تلاش کنید از خروجی‌های مدل استفاده کنید. بعد از مشخص بودن خروجی‌ها، مراکز توپ‌ها را به دست آورید و سپس توپ‌ها را براساس فاصله از توپ سفید رتبه‌بندی کنید. برای پنج تصویر به دلخواه از داده‌ها این رتبه‌بندی‌ها را انجام دهید و نتایج را در گزارش بیاورید. برای این قسمت باید کد فایل با نام `detection.py` را کپی کنید و در یک فایل با نام `detection2.py` بریزید و سپس تغییرات لازم را بر روی آن اعمال کنید. در انتها در کنار گزارش تنها این فایل را به همراه وزن‌های شبکه‌ی به دست آمده آپلود کنید.

سوال ۲ Semantic Segmentation

در این تمرین هدف پیاده سازی شبکه U-Net به منظور انجام Semantic Segmentation برای دیتاست [Cam Vid](#) می باشد. شبکه U-Net شبکه‌ای است که به منظور انجام Semantic Segmentation بر روی داده‌های پزشکی توسعه یافته است ، در این شبکه با استفاده از لایه‌های کانولوشنی و سپس استفاده از Up Sampling ویژگی‌های مربوط به Context عکس را استخراج کرده و با کنار هم قرار دادن این ویژگی‌ها و feature map های استخراج شده در هر مرحله اطلاعات local و Context را ترکیب می کند برای مشاهده مقاله می توانید به این [لینک](#) مراجعه فرمایید.

تذکر: برای قسمت مربوط به Up Sampling از الگوریتم نزدیکترین همسایه و برای تابع خطا نیز از Cross entropy استفاده نمایید.

۱. شبکه U-Net را مطابق آنچه در مقاله ذکر شده پیاده سازی کنید. نمودار خطای داده‌های آموزش و تست را نمایش دهید.

۲. یک عکس از مجموعه داده تست را به شبکه وارد کرده و خروجی آن را به همراه عکس ورودی در گزارش نمایش دهید.

۳. بررسی کنید که تعداد ایپاک لازم برای آموزش شبکه چقدر است، همچنین معیاری که برای جدا سازی داده‌های آموزش و تست استفاده شده را توضیح دهید.

شبکه V-Net ، شبکه‌ی دیگری است که بر روی داده‌های پزشکی توسعه یافته است، با مطالعه مقاله [V-Net](#) ساختار شبکه ، تابع خطای استفاده شده و نتایج را در چند پاراگراف توضیح دهید.

نکات:

- مهلت تحویل این تمرین 10 خرداد است.
- گزارش را در قالب تهیه شده که روی صفحه درس در Elearn بارگذاری شده، بنویسید.
- گزارش شما در فرآیند تصحیح از اهمیت ویژه‌ای برخوردار است. لطفاً تمامی نکات و فرض‌هایی که برای پیاده‌سازی‌ها و محاسبات خود در نظر می‌گیرید را در گزارش ذکر کنید.
- در گزارش خود برای تصاویر زیرنویس و برای جداول هم بالانویس اضافه کنید.
- الزامی به ارائه توضیح جزئیات کد در گزارش نیست. اما باید نتایج به‌دست آمده را گزارش و تحلیل کنید.
- دستیاران آموزشی ملزم به اجرا کردن کدهای شما نیستند. بنابراین هرگونه نتیجه و یا تحلیلی که در شرح سوال از شما خواسته شده است را به طور واضح و کامل در گزارش بیاورید. در صورت عدم رعایت این مورد، بدیهی است که از نمره تمرین کسر می‌شود.
- در صورت مشاهده تقلب امتیاز تمامی افراد شرکت‌کننده در آن، 100- لحاظ می‌شود.
- برای انجام تمرین‌ها و مینی پروژه‌ها، تنها زبان برنامه نویسی مجاز Python است.
- استفاده از کدهای آماده برای تمرین‌ها به هیچ‌وجه مجاز نیست. اما برای مینی پروژه‌ها فقط برای قسمت‌هایی از کد و به عنوان راهنمایی برای پیاده‌سازی، می‌توانید از کدهای آماده استفاده کنید.
- نحوه محاسبه تاخیر به این شکل است: مهلت ارسال بدون جریمه تا تاریخ اعلام شده و پس از آن به مدت هفت روز تا 17 خرداد بارگذاری ممکن است و در نهایت، پس از بازه تاخیر نمره تکلیف صفر خواهد شد. به ازای هر روز تاخیر 5 درصد نمره کاهش پیدا می‌کند.
- لطفاً گزارش، فایل کدها و سایر ضمایم مورد نیاز را با فرمت زیر در سامانه مدیریت دروس بارگذاری نمایید.

EXTRA_HW1_[Lastname]_[StudentNumber].zip

- در صورت وجود هرگونه ابهام یا مشکل می‌توانید از طریق رایانامه‌های زیر با دستیاران آموزشی مربوطه آقایان راستین سورکی (سوال اول) و محمد رضا عبداللهی (سوال دوم) در تماس باشید:

Rastin30@gmail.com

mr.abdollahi@ut.ac.ir