

به نام خدا



دانشگاه تهران



دانشکده مهندسی برق و کامپیوتر

درس شبکه‌های عصبی و یادگیری عمیق

تمرین سوم

کیانا هوشانفر	نام دستیار طراح	پرسش ۱
k.hooshanfar02@gmail.com	رایانامه	
مریم دادخواه	نام دستیار طراح	پرسش ۲
M.dadkhah99@gmail.com	رایانامه	
۱۴۰۳.۰۲.۲۳	مهلت ارسال پاسخ	

فهرست

قوانین	۱
پرسش ۱. پیاده سازی مدل U-Net	۳
۱-۱. آماده سازی مجموعه داده	۳
۲-۱. پیاده سازی مدل	۳
۳-۱. تقویت داده	۳
۴-۱. بهینه ساز، متریک ها و تابع هزینه	۴
۵-۱. آموزش مدل	۴
۶-۱. ارزیابی مدل	۴
پرسش ۲. تشخیص موجودات زیر آب	۵
۱-۲. معرفی مدل Faster R-CNN	۵
۲-۲. سوالات تشریحی	۶
۱-۲-۲. مقایسه مدل های Region-based CNNs	۶
۲-۲-۲. مقایسه مدل های one-stage و two-stage	۶
۳-۲-۲. GIOU, Soft-NMS, OHEM	۶
۳-۲. معرفی مجموعه دادگان	۶
۱-۳-۲. EDA و پیش پردازش دادگان	۷
۲-۳-۲. تقویت داده	۷
۳-۳-۲. ساخت دیتالودر	۸
۴-۲. تعریف مسئله	۸
۱-۴-۲. طراحی معماری Faster R-CNN	۸
۲-۴-۲. Region proposal network طراحی	۹
۲-۴-۲. آموزش مدل	۹
۳-۴-۲. ارزیابی مدل	۹

شکل‌ها

شکل ۱. شمایی کلی از معماری Faster R-CNN ۵

قبل از پاسخ دادن به پرسش‌ها، موارد زیر را با دقت مطالعه نمایید:

- از پاسخ‌های خود یک گزارش در قالبی که در صفحه‌ی درس در سامانه‌ی Elearn با نام **REPORTS_TEMPLATE.docx** قرار داده شده تهیه نمایید.
- پیشنهاد می‌شود تمرین‌ها را در قالب گروه‌های دو نفره انجام دهید. (بیش از دو نفر مجاز نیست و تحویل تک نفره نیز نمره‌ی اضافی ندارد) توجه نمایید الزامی در یکسان ماندن اعضای گروه تا انتهای ترم وجود ندارد. (یعنی، می‌توانید تمرین اول را با شخص A و تمرین دوم را با شخص B و... انجام دهید)
- **کیفیت گزارش شما در فرآیند تصحیح از اهمیت ویژه‌ای برخوردار است؛** بنابراین، لطفا تمامی نکات و فرض‌هایی را که در پیاده‌سازی‌ها و محاسبات خود در نظر می‌گیرید در گزارش ذکر کنید.
- در گزارش خود مطابق با آنچه در قالب نمونه قرار داده شده، برای شکل‌ها زیرنویس و برای جدول‌ها بالانویس در نظر بگیرید.
- الزامی به ارائه توضیح جزئیات کد در گزارش نیست، اما باید نتایج بدست آمده از آن را گزارش و تحلیل کنید.
- **تحلیل نتایج الزامی می‌باشد، حتی اگر در صورت پرسش اشاره‌ای به آن نشده باشد.**
- **دستیاران آموزشی ملزم به اجرا کردن کدهای شما نیستند؛** بنابراین، هرگونه نتیجه و یا تحلیلی که در صورت پرسش از شما خواسته شده را به طور واضح و کامل در گزارش بیاورید. در صورت عدم رعایت این مورد، بدیهی است که از نمره تمرین کسر می‌شود.
- **کدها حتما باید در قالب نوت‌بوک با پسوند .ipynb تهیه شوند، در پایان کار، تمامی کد اجرا شود و خروجی هر سلول حتما در این فایل ارسالی شما ذخیره شده باشد.** بنابراین برای مثال اگر خروجی سلولی یک نمودار است که در گزارش آورده‌اید، این نمودار باید هم در گزارش هم در نوت‌بوک کدها وجود داشته باشد.
- **در صورت مشاهده‌ی تقلب امتیاز تمامی افراد شرکت‌کننده در آن، 100- لحاظ می‌شود.**
- تنها زبان برنامه نویسی مجاز **Python** است.
- **استفاده از کدهای آماده برای تمرین‌ها به هیچ وجه مجاز نیست.** در صورتی که دو گروه از یک منبع مشترک استفاده کنند و کدهای مشابه تحویل دهند، تقلب محسوب می‌شود.
- نحوه محاسبه تاخیر به این شکل است: پس از پایان رسیدن مهلت ارسال گزارش، حداکثر تا یک هفته امکان ارسال با تاخیر وجود دارد، پس از این یک هفته نمره آن تکلیف برای شما صفر خواهد شد.

○ سه روز اول: بدون جریمه

○ روز چهارم: ۵ درصد

○ روز پنجم: ۱۰ درصد

○ روز ششم: ۱۵ درصد

○ روز هفتم: ۲۰ درصد

- حداکثر نمره‌ای که برای هر سوال می‌توان اخذ کرد ۱۰۰ بوده و اگر مجموع بارم یک سوال بیشتر از ۱۰۰ باشد، در صورت اخذ نمره بیشتر از ۱۰۰، اعمال نخواهد شد.

○ برای مثال: اگر نمره اخذ شده از سوال ۱ برابر ۱۰۵ و نمره سوال ۲ برابر ۹۵ باشد، نمره نهایی تمرین ۹۷.۵ خواهد بود و نه ۱۰۰.

- لطفا گزارش، کدها و سایر ضمایم را به در یک پوشه با نام زیر قرار داده و آن را فشرده سازید، سپس در سامانه‌ی Elearn بارگذاری نمایید:

HW [Number] _ [Lastname] _ [StudentNumber] _ [Lastname] _ [StudentNumber].zip

(مثال: HW1_Ahmadi_810199101_Bagheri_810199102.zip)

- برای گروه‌های دو نفره، بارگذاری تمرین از جانب یکی از اعضا کافی است ولی پیشنهاد می‌شود هر دو نفر بارگذاری نمایند.

پرسش ۱. پیاده سازی مدل U-Net

در این پرسش، شما بر روی پیاده سازی مدل U-Net برای مسئله Semantic Segmentation کار خواهید کرد. تمرکز اصلی سوال، درک مفاهیم Segmentation تصویر و کار با مدل معرفی شده است.

۱-۱. آماده سازی مجموعه داده

(۵ نمره)

مجموعه داده انتخاب شده برای این پرسش، [تصاویر MRI مغز](#) می باشد. دادگان را دانلود کرده، تصاویر را خوانده و نمونه‌ای از آن‌ها را (تصویر اصلی به همراه ماسک) نمایش دهید. سپس داده‌ها را به نسبت ۸ - ۱ - ۱ به دو بخش آموزش^۱، اعتبارسنجی^۲ و ارزیابی^۳ تقسیم کنید.

۲-۱. پیاده‌سازی مدل

(۲۵ نمره)

برای پیاده سازی مدل از [این مقاله](#) استفاده کنید. [مقاله U-Net](#) در سال ۲۰۱۵ ارائه شد، در مورد نحوه کار مدل توضیح داده و ساختار معماری مدل انتخاب شده را شرح دهید.

۳-۱. تقویت داده

(۱۵ نمره)

تکنیک‌های تقویت داده^۴ شامل اعمال تغییراتی بر روی تصاویر از طریق روش‌های مختلفی است تا به این ترتیب، قابلیت تعمیم‌دهی و بخش‌بندی بهتر تصاویر توسط مدل در شرایط متفاوت بهبود یابد. بسته به

¹ Train

² Validation

³ Test

⁴ Data Augmentation

ماهیت داده‌ها، روش‌های مناسب برای تقویت انتخاب می‌شوند. در مورد دلایل انتخاب هر یک از این روش‌ها و تأثیرگذاریشان توضیح دهید. پس از پیاده‌سازی، چند نمونه از تصاویر تقویت شده را نمایش دهید. در این قسمت می‌توانید از کتابخانه [Albumentations](#) استفاده کنید.

۴-۱. بهینه ساز، متریک‌ها و تابع هزینه

(۲۰ نمره)

دو متریک Dice Coefficient و IoU Score را خودتان پیاده سازی کرده و در مورد هر یک توضیحاتی ارائه دهید. همچنین از آنها به عنوان متریک برای آموزش شبکه استفاده کنید. بهینه ساز و تابع هزینه را مطابق مقاله اول پیاده سازی کنید.

(متریک‌هایی که باید گزارش دهید: Accuracy, MeanIoU, Dice coefficient)

۵-۱. آموزش مدل

(۲۰ نمره)

از دادگان آموزشی برای آموزش مدل و از دادگان ارزیابی جهت بررسی کیفیت آموزش مدل (با گزارش مقادیر تابع هزینه و متریک) استفاده کنید و مدلی که در قسمت قبل پیاده سازی کرده‌اید را به میزان کافی آموزش دهید (تعداد epochها و batch size به اختیار شماست). در انتهای آموزش مدل نمودار تغییرات متریک‌ها و تابع هزینه در طول آموزش را نیز گزارش کنید. در این قسمت می‌توانید از early-stopping استفاده کنید.

۶-۱. ارزیابی مدل

(۱۵ نمره)

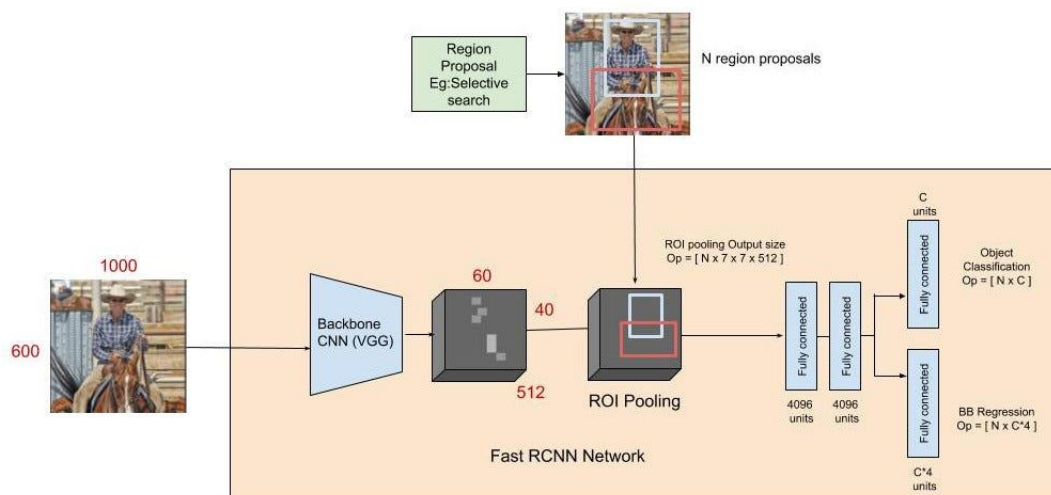
بعد از آموزش مدل به میزان کافی، ۵ نمونه از تصاویر ارزیابی را به همراه ماسک واقعی، ماسک پیش‌بینی شده نشان دهید. تحلیل خود را از نتایج بدست آمده روی تصاویر و مقادیر محاسبه شده تابع هزینه و متریک‌ها در طی آموزش روی دادگان آموزشی و ارزیابی بیان کنید.

پرسش ۲. تشخیص موجودات زیر آب

در این پرسش با کاربرد معماری Faster R-CNN در بحث تشخیص اشیاء^۱، آشنا خواهید شد و مقاله‌ای را در این رابطه پیاده‌سازی خواهید کرد. مقاله‌ای که برای استفاده در این سوال مورد نظر است با عنوان "Underwater Object Detection Method Based on Improved Faster RCNN" در پیوست ارسال شده است. در ادامه قصد داریم از برخی از ایده‌های به کار رفته در این مقاله استفاده کرده و آن‌ها را پیاده سازی کنیم.

۱-۲. معرفی مدل Faster R-CNN

معماری Faster R-CNN یک مدل شبکه عصبی معروف در بحث تشخیص اشیاء است که با استفاده از مفهوم region-based CNNs، یک چهارچوب یکپارچه برای تشخیص اشیاء ارائه می‌دهد. این مدل جزو دسته معماری‌های دو مرحله‌ای^۲ بوده و شامل یک شبکه Region Proposal Network (RPN) و یک شبکه تشخیص^۳ است. دسترسی به مقاله Faster R-CNN از طریق [این لینک](#) امکان پذیر است.



شکل ۱. شمایی کلی از معماری Faster R-CNN

¹ Object detection

² Two-stage

³ Detection network

۲-۲. سوالات تشریحی

(۱۰ نمره)

۱-۲-۲. مقایسه مدل‌های Region-based CNNs

سه مدل معروفی که در بحث تشخیص اشیاء با استفاده از مدل‌های Region-based CNNs وجود دارند، R-CNN، Fast R-CNN و Faster R-CNN می‌باشند. درباره معماری هر یک مختصراً توضیح داده و بیان کنید هر کدام چه مزایایی نسبت به دیگری داشته و چه مشکلی را برطرف می‌کنند.

۲-۲-۲. مقایسه مدل‌های one-stage و two-stage

معماری‌هایی که در تشخیص اشیاء وجود دارند به دو دسته one-stage detector و two-stage detector دسته‌بندی می‌شوند. در خصوص مزایا و معایب هر کدام از این دو دسته معماری تحقیق کرده و کاربردها و نمونه‌های معروف هر کدام را ذکر کنید.

۳-۲-۲. GIOU, Soft-NMS, OHEM

سه مورد از مفاهیم جدیدی که در مقاله^۱ استفاده شده است، OHEM به عنوان تکنیکی برای بالانس کردن نمونه پروپوزال‌های مثبت و منفی، soft-NMS در مقابل NMS و GIOU در مقابل IOU است. هر کدام را مختصراً توضیح داده و مزایا و کاربردهای آن را ذکر کنید.

۳-۲. معرفی مجموعه دادگان

مجموعه داده مورد استفاده برای این بخش Underwater Object Detection نام دارد که از طریق [kaggle](https://www.kaggle.com/datasets/underwater-object-detection) در دسترس است. این مجموعه دادگان شامل ۶۳۸ تصویر از موجودات زیر آب است که در ۷ کلاس طبقه‌بندی می‌شوند. همچنین دیتاست شامل موقعیت جعبه‌های مرزی^۲ برای هر موجود داخل تصویر است که آن را برای استفاده در تسک تشخیص اشیاء مناسب می‌کند.

^۱ Underwater Object Detection Method Based on Improved Faster RCNN

^۲ Bounding boxes

۱-۳-۲. EDA^۱ و پیش پردازش دادگان

(۱۰ نمره)

در این قسمت لازم است برای داشتن دید بهتر نسبت به مجموعه دادگان مورد استفاده، آنها رو مورد ارزیابی قرار دهید. برای بخش EDA لازم است مواردی همچون بررسی آماری تعداد اشیاء داخل تصاویر، تعداد تصاویر در هر کلاس و غیره را بررسی کرده و سپس پیش پردازش‌های لازم برای انجام تسک مورد نظر را بیان کنید. علت استفاده از هر روشی برای EDA و یا پیش پردازش دادگان را در گزارش خود ذکر کنید. برای مجموعه دادگان train، validation و test از داده‌های با همین اسامی در دیتاست استفاده کنید.

۲-۳-۲. تقویت داده^۲

(۱۰ نمره)

روش‌های تقویت داده، تغییراتی را در تصاویر از طریق انجام تکنیک‌هایی ارائه می‌کنند تا بدین ترتیب، توانایی مدل را در تعمیم دهی بهتر در شرایط مختلف افزایش دهند. با توجه به ماهیت داده‌ها، روش‌های مناسبی را برای تقویت داده انتخاب کرده، علت انتخاب و نحوه تاثیرگذاری هر یک را توضیح داده و بعد از پیاده سازی، چند نمونه تقویت شده را نمایش دهید.

(۵ نمره امتیازی)

یکی از تکنیک‌های بکار رفته در مقاله برای تقویت دادگان، Mosaic Augmentation است. درباره این شیوه افزایش داده توضیح داده و مزایا، معایب و کاربردهای آن را ذکر کنید. همچنین این تکنیک را برای افزایش دادگان دیتاست خود پیاده سازی کنید.

^۱ Exploratory Data Analysis

^۲ Data Augmentation

۳-۳-۲. ساخت دیتالودر

(۱۰ نمره)

همانطور که از قبل می‌دانید طول همه دنباله‌هایی که در یک Batch قرار می‌گیرند باید باهم برابر باشد. از آنجایی که تعداد اشیاء موجود در تصاویر مختلف با هم متفاوت است، بنابراین برای هر تصویر داخل batch تعداد متفاوتی bounding-box position و object-label خواهیم داشت. لذا لازم است دیتالودر را طوری پیاده سازی نمایید که طول همه داده‌های درون یک Batch به اندازه‌ی طول بلندترین دنباله موجود در آن Batch باشد.

۴-۲. تعریف مسئله

در این سوال قرار است تا به مسئله تشخیص موجودات زیر آب با استفاده از مدل‌های مبتنی بر region-based CNNs بپردازیم. برای این کار لازم است یک مدل Faster R-CNN را از ابتدا پیاده سازی کرده و از آن برای تشخیص اشیاء استفاده کنید.

۴-۱. طراحی معماری Faster R-CNN

(۱۵ نمره)

برای طراحی معماری شبکه Faster R-CNN نیاز داریم که از برخی از ایده‌هایی که در مقاله پیوست شده مطرح شده است، استفاده کنیم. برای مدل استخراج ویژگی^۱ در معماری مدل اصلی، از مدل از پیش آموزش دیده resnet101^۲ استفاده کنید. می‌توانید مدل را از این لینک بارگذاری کنید. برای سایر بخش‌های معماری مدل، از معماری که در مقاله اصلی Faster R-CNN آورده شده است استفاده کنید ولی تنها تغییری که لازم است در این معماری وارد کنید، استفاده از تکنیک GIOU به جای IOU است که در مقاله پیوست شده، آورده شده است.

^۱ Feature extractor

^۲ Pre-trained

۲-۴-۲. طراحی Region proposal network

(۱۵ نمره)

یکی از بخش‌های مهم در معماری مدل‌های Two-stage detector شبکه ایجاد region proposal است که در این قسمت می‌خواهیم این شبکه را پیاده سازی کنیم. توجه داشته باشید که در این شبکه لازم است تعدادی Anchor point روی تصویر ورودی ایجاد کرده و برای هر کدام k عدد Anchor box با سائزهای متفاوت، ایجاد کنید. Anchor points در واقع نقاطی روی Feature map هستند که Box های پیشنهادی اولیه روی آن‌ها ایجاد می‌شوند. تابعی برای ایجاد این نقاط و anchor box های آن پیاده سازی کرده و از آنها روی ۲ عدد از تصاویر مجموعه دادگان استفاده کرده و آن‌ها را نمایش دهید. (الگوریتم انتخاب مقدار k را خودتان تعیین کرده و علت انتخاب خود را توضیح دهید. می‌توانید از ایده مطرح شده در مقاله پیوست شده استفاده کنید.) بعد از ایجاد Anchor boxes، با استفاده از مقادیر واقعی آن‌ها در مجموعه دادگان، نمونه‌های Positive و Negative را با مقدار Threshold مناسب، استخراج کرده و آن‌ها را برای ۲ نمونه از تصاویر مجموعه دادگان نمایش دهید. مقادیر پارامترهایی که در این قسمت استفاده می‌کنید را در گزارش خود ذکر کنید.

۲-۴-۲. آموزش مدل

(۱۵ نمره)

در این بخش می‌خواهیم با استفاده از مدلی که در بخش قبل پیاده سازی کردید، یک مدل کامل را برای تشخیص اشیاء داخل تصویر آموزش دهیم. درباره تابع loss مورد استفاده در آموزش مدل‌ها تحقیق کرده و آن را پیاده سازی کنید. توجه کنید که گزارش کامل نتایج، مقادیر پارامترها و نمودار روند آموزش در این بخش اهمیت دارد. برای آموزش از دادگان آموزش، اعتبار سنجی و ارزیابی به نحوی که در پوشه اصلی دادگان وجود دارد استفاده کنید.

۲-۴-۳. ارزیابی مدل

(۱۵ نمره)

در این قسمت از مدل آموزش داده شده در قسمت قبل استفاده کرده و ۱۰ نمونه از دادگان تست را به آن بدهید و نمونه تصویر را به همراه جعبه‌های مرزی پیش‌بینی شده توسط مدل و جعبه‌های واقعی آن نمایش دهید.