



به نام خدا  
دانشگاه تهران  
دانشکده مهندسی برق و کامپیوتر



## درس شبکه‌های عصبی و یادگیری عمیق

### تمرین سوم

پارمیس بطحائیان	نام دستیار طراح	پرسش ۱
parmisbathaeiyan@gmail.com	رایانامه	
رضا براتی	نام دستیار طراح	پرسش ۲
Miladbarati2000@gmail.com	رایانامه	

## فهرست

1	قوانین
3	پرسش 1. سگمنتیشن تصاویر شهری
3	1-1. توصیف مدل ارائه شده
3	1-2. آماده سازی مجموعه داده
3	1-3. بهینه ساز، متریک ها و تابع هزینه
4	1-4. پیاده سازی مدل
4	1-5. آموزش مدل
4	1-6. ارزیابی مدل
5	پرسش 2 : Oriented R-CNN برای تشخیص اشیاء
5	بخش اول: سوالات نظری (50 نمره)
5	درک مفهومی (10 نمره)
5	اجزای مدل (15 نمره)
5	Rotated RoI Align (10 نمره)
5	عملکرد و کارایی (15 نمره)
6	بخش دوم: پیاده سازی عملی (60 نمره)
6	راه اندازی محیط و آماده سازی مجموعه داده (15 نمره)
6	آموزش مدل Oriented R-CNN (20 نمره)
6	ارزیابی و تحلیل نتایج (15 نمره)
7	تحلیل مقایسه ای و پیشنهاد بهبود (10 نمره)

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 نمره 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. سگمنتیشن تصاویر شهری

در این تمرین، شما بر روی پیاده‌سازی مدل Fast SCNN برای مسئله‌ی سگمنتیشن تصاویر شهری کار خواهید کرد. هدف این سوال، درک اصول سگمنتیشن تصویر و به‌کارگیری مدل معرفی‌شده در [مقاله‌ی پیوست شده](#) است.

### ۱-۱. توصیف مدل ارائه شده

(۱۵ نمره)

مقاله را خوانده و در مورد نحوه‌ی عملکرد مدل پیشنهاد شده، ساختار معماری آن، و نقش هر قسمت در فرآیند سگمنتیشن، توضیح دهید. این مدل را با مدل‌های انکدر دیکدر مانند U-net از لحاظ ساختار و عملکرد مقایسه کنید.

### ۲-۱. آماده‌سازی مجموعه داده

(۵ نمره)

مجموعه داده انتخاب‌شده برای این تمرین CamVid می‌باشد. این مجموعه داده را از [این صفحه](#) [گیت‌هاب](#) دریافت کنید. تصاویر را خوانده و چند نمونه از تصاویر اصلی را همراه با ماسک نمایش دهید. تعداد داده دسته‌های آموزش و اعتبارسنجی را گزارش کنید. دقت داشته باشید که استفاده از روش‌های data augmentation در این تمرین اجباری نیست.

### ۳-۱. بهینه‌ساز، متریک‌ها و تابع هزینه

(۱۵ نمره)

در مورد دو متریک Dice Coefficient و IoU Score توضیح دهید. این دو را خودتان پیاده‌سازی کرده و از آنها به همراه Accuracy به عنوان متریک حین آموزش شبکه استفاده کنید. بهینه‌ساز و تابع هزینه را مطابق با مقاله یا با انتخاب خودتان تنظیم کنید.

## ۴-۱. پیاده‌سازی مدل

(۲۰ نمره)

مدل Fast-SCNN را مطابق مقاله برای مساله سگمنتیشن تصاویر پیاده‌سازی کنید. همچنین تعداد کل پارامترهای مدل ایجاد شده را گزارش کنید. در بخش گزارش خود، در مورد ساختار بلاک‌های اصلی استفاده شده در مدل معرفی شده، یعنی **Depthwise Separable Convolution**، **Depthwise Convolution**، **Inverted Residual Block** و **Pyramid Pooling Module** شرح مختصری ارائه دهید.

## ۵-۱. آموزش مدل

(۳۰ نمره)

برای آموزش مدل تعداد epoch ها و batch size را به دلخواه انتخاب کنید و مقدار آنها را همراه مقدار هایپرپارامترهای دیگر، شامل تابع هزینه، optimizer و learning rate گزارش کنید. در پایان، نمودار تابع هزینه، دقت، IoU Score و Dice Coefficient را روی داده های آموزش و ارزیابی را رسم کرده و تحلیل کنید. توجه داشته باشید که رسیدن به نتیجه مقاله لزومی ندارد و رسیدن متریک های IoU Score و Dice Coefficient به مقدار 0.5 روی داده های اعتبارسنجی، نشان دهنده یادگیری خوب مدل شماست.

## ۶-۱. ارزیابی مدل

(۱۵ نمره)

پس از آموزش مدل، ۱۰ نمونه از تصاویر ارزیابی را همراه با ماسک واقعی و ماسک پیش‌بینی شده نمایش دهید. با توجه به نتایج به‌دست‌آمده روی تصاویر، عملکرد مدل را توضیح دهید.

## پرسش ۲: Oriented R-CNN برای تشخیص اشیاء

این تمرین به بررسی مدل [Oriented R-CNN](#) برای تشخیص اشیاء می‌پردازد که در مقاله "Oriented R-CNN for Object Detection" ارائه شده است. این تمرین به منظور تقویت درک نظری شما از روش‌های پیشرفته تشخیص اشیاء جهت‌دار و ارائه تجربه عملی در پیاده‌سازی و ارزیابی مدل‌های یادگیری عمیق طراحی شده است. پرسش‌های نظری میزان درک شما از مفاهیم، معماری‌ها و تکنیک‌های مطرح‌شده در مقاله را می‌سنجد، در حالی که وظایف عملی شما را در راه‌اندازی محیط، آموزش، ارزیابی و تحلیل عملکرد چارچوب Oriented R-CNN هدایت می‌کند.

### بخش اول: سوالات نظری (۵۰ نمره)

#### درک مفهومی (۱۰ نمره)

الف. انگیزه اصلی توسعه Oriented R-CNN را توضیح دهید. این مدل چه محدودیت‌هایی از روش‌های قبلی را برطرف می‌کند؟ مثال‌هایی ارائه کنید. (۵ نمره)

ب. مزایای استفاده از نمایش "midpoint offset" نسبت به نمایش‌های سنتی جعبه‌های محدودکننده را توضیح دهید. مثال‌هایی برای روشن‌شدن توضیحات خود ارائه کنید. (۵ نمره)

#### اجزای مدل (۱۵ نمره)

الف. معماری Oriented RPN را شرح دهید و تفاوت آن را با RPN سنتی توضیح دهید. برای روشن‌تر شدن توضیحات خود از نمودار یا طرح استفاده کنید. (۷ نمره)

ب. نحوه فرمول‌بندی تابع هزینه (loss function) در Oriented RPN را توضیح دهید. هر یک از اجزای این تابع را به‌طور واضح تعریف و هدف آن‌ها را شرح دهید. (۸ نمره)

#### Rotated RoI Align (۱۰ نمره)

الف. هدف از Rotated RoIAlign چیست؟ گام‌به‌گام توضیح دهید که چگونه این عملیات انجام می‌شود و یک مثال روشن ارائه کنید. (۵ نمره)

ب. مشکلات احتمالی در صورت عدم استفاده از Rotated RoIAlign را توضیح دهید. برای استدلال خود مثال‌ها یا توجیه‌های نظری ارائه کنید. (۵ نمره)

#### عملکرد و کارایی (۱۵ نمره)

الف. توضیح دهید که Oriented R-CNN چگونه به دقت و کارایی بالا دست پیدا می‌کند. به‌طور مشخص به آزمایش‌ها و نتایج ارائه‌شده در مقاله ارجاع دهید. (۵ نمره)

ب. عوامل مؤثر در کارایی محاسباتی چارچوب Oriented R-CNN را توضیح دهید و نقش هر یک از این عوامل را مشخص کنید. (۵ نمره)

ج. یک تحلیل انتقادی کوتاه از مقایسه Oriented R-CNN با سایر آشکارسازهای جهت‌دار دو مرحله‌ای ذکر شده در مقاله ارائه کنید و نقاط قوت و ضعف کلیدی را برجسته کنید. (۵ نمره)

### بخش دوم: پیاده‌سازی عملی (۶۰ نمره)

#### راه‌اندازی محیط و آماده‌سازی مجموعه داده (۱۵ نمره)

مجموعه داده HRSC2016 را با استفاده از کد ارائه شده دانلود و پیش‌پردازش کنید. یک کلاس سفارشی PyTorch Dataset پیاده‌سازی کنید که داده‌ها را بارگذاری و حاشیه‌نویسی‌ها را به نمایش-midpoint-offset تبدیل می‌کند. قطعاتی از کد و حداقل سه مثال از بارگذاری موفق داده‌ها ارائه دهید. (۱۵ نمره)

```
import kagglehub
```

```
dataset_path = kagglehub.dataset_download('weiming97/hrsc2016-ms-dataset')
```

#### آموزش مدل Oriented R-CNN (۲۰ نمره)

الف. مدل Oriented R-CNN را (بدون استفاده از MMDetection) با استفاده از ResNet-50 و FPN مطابق مقاله به مدت ۳۶ دوره (epoch) آموزش دهید. تنظیمات آموزش خود شامل برنامه نرخ یادگیری، تنظیمات بهینه‌ساز، اندازه دسته‌ها و هرگونه افزایش داده‌ها (augmentation) را به‌طور واضح مستند کنید. (۱۰ نمره)

ب. نمودارهای مربوط به توابع هزینه در فرآیند آموزش را رسم و تحلیل کنید. تفسیر این نمودارها را ارائه دهید و الگوهای همگرایی، علائم بیش‌برازش و پایداری کلی آموزش را بررسی کنید. (۱۰ نمره)

#### ارزیابی و تحلیل نتایج (۱۵ نمره)

الف. مدل آموزش‌دیده خود را با نمایش پروپوزال‌ها و جعبه‌های Ground Truth روی تصاویر مجموعه آزمون ارزیابی کنید. حداقل پنج تصویر با پیش‌بینی‌های مدل در مقایسه با Ground Truth ارائه دهید. تحلیل دقیق هر تصویر شامل دقت پیش‌بینی‌ها، خطاها و دلایل احتمالی پیش‌بینی‌های نادرست را شرح دهید. (۱۵ نمره)



### تحلیل مقایسه‌ای و پیشنهاد بهبود (۱۰ نمره)

الف. عملکرد مدل خود را با معیارهای ارائه‌شده در مقاله اصلی مقایسه کنید. نتایج خود را به‌طور واضح بیان کرده و تفاوت‌های احتمالی با مقاله را تحلیل کنید. به دلایل احتمالی این تفاوت‌ها اشاره و پیشنهادهای مشخص و عملی برای بهبود پیاده‌سازی فعلی خود ارائه دهید. (۱۰ نمره)