|  |  |  |
| --- | --- | --- |
|  | **تمرین سری ششم درس مبانی بینایی کامپیوتر** | **نام مدرس:** دکتر محمدرضا محمدی  **دستیاران آموزشی مرتبط**:علیرضا خیرخواه، امیرحسین حسینی،کسری شریعتی، ،آروین ستاری |
| **مهلت تحویل )بدون کسر نمره:( سه شنبه 72 خرداد** |

# بخش تئوری

1. در ارتباط با object detection به سوالات زیر پاسخ دهید:(21نمره)

الف) برای پیدا کردن پروپوزال‌ها در بعضی مواقع معیار IOU خوب جواب نمیدهد(کیفیت لازم را ندارد)، آن موارد که IOU دچار خطا

میشود را نام برده و معیارهای مکمل یا جایگزین برای ارزیابی بهتر را بنویسید.

مقاله پیشنهادی: [**Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression]**

ب) Proposals with high recall but at low overlap are not effective for detection. این جمله را تحلیل کرده و با مثال

توضیح دهید.

مقاله پیشنهادی: [**What makes for effective detection proposals?]**

پ) تفاوت anchor boxes در SSD با سایر مدل های detection مانند Faster R-CNN چیست؟

ت) چرا در SSD نیازی به استفاده از region proposal methods مانند Selective Search یا RPN وجود ندارد؟

ث) در مقاله SSD ، از hard negative mining استفاده شده است. این تکنیک چیست و چرا برای آموزش پایدار شبکه مهم است؟

مقاله پیشنهادی: [**SSD: Single Shot MultiBox Detector]**

ج) در مقاله "SSD"، نویسندگان عملکرد SSD را از نظر میانگین دقت (mAP) و نرخ فریم بر ثانیه (FPS) با مدل‌های Faster R-CNN و YOLOv1 مقایسه کرده‌ اند. با استناد به اعداد و نتایج ارائه‌ شده در مقاله، مشخص کنید:

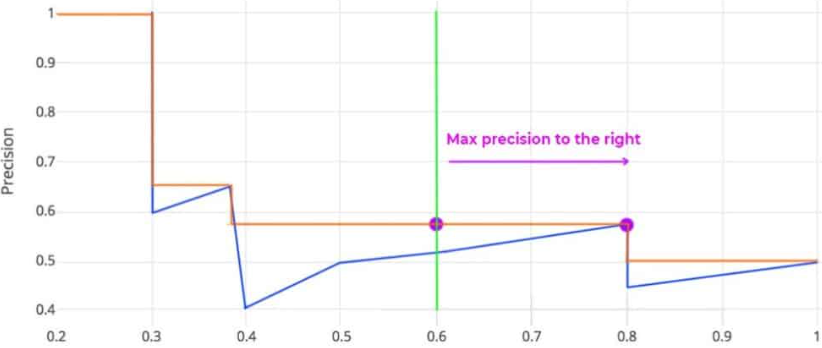
* SSD از نظر دقت و سرعت در چه جایگاهی قرار دارد؟
* چه (trade-off) ای میان سرعت و دقت ارائه می‌ دهد؟
* برای چه کاربردهایی SSD نسبت به سایر روش‌ها انتخاب بهتری است؟

مقاله پیشنهادی: [**SSD: Single Shot MultiBox Detector]**

1. در نمودار مختصاتی زیر که محور افقی نشان‌دهنده Recall و محور عمودی Precision است، دو منحنی رسم شده اند.  
    یکی از این منحنی ها نمایانگر مقادیر واقعی Precision (Actual Precision) و دیگری نمایانگرPrecision درونیابی‌شده

(Interpolated Precision) است که در محاسبه‌ی Average Precision به کار می رود. مشخص کنید کدام منحنی actual و کدام interpolated است، و علت انتخاب خود را تحلیل کنید. چرا Interpolated Precision لازم است؟ هرکدام چه کاربردی را

پوشش میدهد؟ همچنین، نحوه‌ی محاسبه ی Average Precision براساس این نمودار را توضیح دهید.(8نمره)



1

|  |  |  |
| --- | --- | --- |
|  | **تمرین سری ششم درس مبانی بینایی کامپیوتر** | **نام مدرس:** دکتر محمدرضا محمدی  **دستیاران آموزشی مرتبط**:علیرضا خیرخواه، امیرحسین حسینی،کسری شریعتی، آروین ستاری |
| **مهلت تحویل )بدون کسر نمره:( سه شنبه 72 خرداد** |

1. سوالات زیر در ارتباط با Semantic Segmentation پاسخ دهید(12نمره)

الف) منظور از Fully Convolutional در FCN چیست؟ چه تفاوتی با معماری‌ های CNN معمولی برای classification دارد؟

چرا در معماری FCN، لایه‌های Fully Connected حذف شده‌ اند؟ این حذف چه مزایا و چالش‌ هایی ایجاد می‌ کند؟

ب) FCN از pre-trained classification models مثل AlexNet, VGG, GoogLeNet استفاده می‌ کند.

چرا استفاده از transfer learning در segmentation کاربردی است؟

ج) چرا pooling در CNNها باعث از دست رفتن دقت مکانی می‌شود و FCN برای حل آن چه راه‌ حلی پیشنهاد می‌ کند؟

1. به سوالات زیر در ارتباط با Focal Loss پاسخ دهید:(11نمره)

الف) هنگام استفاده از Focal Loss برای مقابله با class imbalance، این تکنیک چه تاثیری بر precision و recall دارد؟ از نظر

مفهومی آیا میتوان گفت focal loss نسخه ی ریگولاریزه شده ی cross entropy است؟ اگر γ از صفر تا بی نهایت تغییر کند،

چه اتفاقی می افتد؟ دقیق تحلیل کنید.

ب) اگر Focal Loss از RetinaNet حذف شود و به جای آن از Cross-Entropy معمولی استفاده شود، چه پدیده ‌ای در training دیده میشود؟

ج) فرض کنید بخواهید RetinaNet را روی یک داده با 100 کلاس آموزش دهید که تنها 10 کلاس آن متداول ‌اند. چه تغییری در

focal loss یا تنظیمات anchor پیشنهاد می ‌دهید؟

1. در الگوریتم SORT (Simple Online and Realtime Tracking)، از الگوریتم Hungarian برای به حداقل رساندن هزینه کلی  
    تخصیص n تشخیص (detections) به m مسیر ردیابی (tracks) استفاده می‌شود. این تخصیص بر اساس   
   یک ماتریس هزینه صورت می‌گیرد.  
   الف) با توجه به ماتریس هزینه زیر، تخصیص بهینه را که هزینه کل را حداقل می‌کند، پیدا کنید.(12نمره)  
   ب) چرا ماتریس هزینه در عمل باید هم فاصله IoU و هم خطای پیش‌بینی حرکت را ترکیب کند؟  
   پ) شبکه سیامی را در نظر بگیرید که تصاویر ورودی را به یک بردار بازنمایی (Embedding) در فضای نگاشت می‌کند و از   
   یک تابع فاصله استفاده می‌کند.  
   اولا) نشان دهید که اگر یک تابع خطی باشد یعنی آنگاه تابع فاصله به یک فرم درجه دو تبدیل می‌شود.

ثانیا) فرض کنید امتیاز شباهت باشد. وقتی γ→∞ چه اتفاقی برای این امتیاز می‌افتد؟   
 این چه معنایی برای اطمینان مدل دارد؟

2

|  |
| --- |
|  |
|  |  |

|  |  |  |
| --- | --- | --- |
|  | **تمرین سری ششم درس مبانی بینایی کامپیوتر** | **نام مدرس:** دکتر محمدرضا محمدی  **دستیاران آموزشی مرتبط**:علیرضا خیرخواه، امیرحسین حسینی،کسری شریعتی، آروین ستاری |
| **مهلت تحویل )بدون کسر نمره:( سه شنبه 72 خرداد** |

6. ساختار زیر یک شبکه‌ی ساده‌ی کانولوشنی برای طبقه‌بندی تصویر با اندازه‌ی ثابت ۳۲×۳۲ است. این شبکه شامل یک لایه‌ی کانولوشنی و دو لایه‌ی Fully Connected است:

model = tf.keras.Sequential([

tf.keras.layers.Conv2D(16, (3, 3), activation='relu', input\_shape=(32, 32, 3)),

tf.keras.layers.Flatten(),

tf.keras.layers.Dense(128, activation='relu'),

tf.keras.layers.Dense(10, activation='softmax')

])

الف) کد هم‌ارز این شبکه را به صورت یک شبکه‌ی تماماً کانولوشنی بنویسید به‌گونه‌ای که از لایه‌های Dense و Flatten استفاده نشود   
و رفتار کلی مدل حفظ شود.

ب) تعداد کل پارامترهای مدل را در هر دو نسخه (اصلی و تماماً کانولوشنی) محاسبه و با هم مقایسه کنید.

ج) با وجود برابر بودن تعداد پارامترها، دست‌کم دو دلیل برای استفاده از شبکه‌های تماماً کانولوشنی در مسائل semantic segmentation بیان کنید.

# بخش عملی

# به نوت‌بوک semantic segmentation مراجعه کنید و بخش‌های خواسته شده را کامل کنید.

# همچنین callbackهای مناسب جهت ذخیره بهترین مدل را به کد اضافه کنید. در این سوال ابتدا یک شبکه Unet را از صفر آموزش می‌دهید.

# سپس در بخش دوم یک شبکه Unet با انکودر پیش‌آموخته را آموزش می‌دهید.

# در مرحله اول انکودر کاملاً فریز است و تنها دیکودر آموزش می‌بیند و در مرحله نهایی کل شبکه تنظیم دقیق می‌شود.

# در گزارش خود علاوه بر توضیح کد خودتان، توضیحات مختصری در مورد قسمت‌های آماده کد ارائه دهید.

# دو سوال در دو نوتبوک 8\_1 و 8\_2 در مورد تعریف anchor boxes برای یک feature map ساده و همچنین اعمال اعمال NMS (Non-Maximum Suppression)

# روی چند جعبه فرضی تهیه و پیوست شده است.

1. دو سوال در نوتبوک در مورد ردیابی اشیاء و همچنین تشخیص صورت تهیه و پیوست شده است.(25نمره)

۳