

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



یادگیری عمیق پیشرفته

پروژه سوم

خرداد ۱۴۰۰

فهرست مسائل

۳Unsupervised Representation Learning - سوال اول
۳بخش اول
۳محاسبه دقت و گزارش
۳بخش دوم
۳بخش سوم
۴SimCLR - سوال دوم
۴بخش اول
۴بخش دوم
۴بخش سوم
۵Contrastive Predictive Coding - سوال سوم
۶Downstream task - سوال چهارم
۶بخش اول
۶بخش دوم

سوال اول – Unsupervised Representation Learning

بخش اول

در سوال اول با استفاده از روش rotation prediction قصد داریم که یک representation مفید از عکس‌ها به دست بیاوریم. به این منظور اول مقاله زیر را بخوانید.

[Unsupervised Representation Learning by Predicting Image Rotations](#)

سپس این روش را برای Cifar10 با استفاده از ResNet18 (که البته stem آن برای cifar عوض شده، به این معنا که stride اولین لایه یک است و max pooling ندارد) پیاده کنید و برای تعداد صد epoch آموزش دهید.

محاسبه دقت و گزارش

پس اتمام آموزش، از روش معمول linear classifier بر روی feature extractor فریز شده، دقت مدل را حساب کنید و این کار را بر روی خروجی stage های مختلف ResNet انجام دهید (برای این کار از global average pooling استفاده کنید).

بخش دوم

روش cut mix که یک data augmentation است را به مدل خود اضافه کنید. برای این کار دو راه برای پیاده سازی دارید، یا اول cut mix را اعمال کنید و بعد دوران یا بالعکس. شما باید این دو روش را با هم مقایسه کنید و در گزارش خود نتیجه را ارائه دهید.

بخش سوم

بجای این که هر چهار دوران مختلف یک عکس را در هر minibatch قرار دهید، از چهار عکس مختلف استفاده کنید و دقت‌ها را مقایسه کنید.

تفاوت نتایج با آنچه در بخش های قبلی به دست آمده را توضیح دهید.

سوال دوم - SimCLR

بخش اول

در این سوال، باید روش SimCLR را بر روی CIFAR10 (مجدداً با استفاده از ResNet18) پیاده کنید. دقت کنید که باید از بیشترین batch size ای که ماشین تان اجازه می دهد استفاده کنید که بهترین نتیجه را بگیرید.

آموزش را برای تعداد صد epoch و تنها با evaluate کردن با استفاده از یک linear classifier روی خروجی stage آخر ResNet انجام دهید و نتیجه را گزارش کنید.

بخش دوم

روش تغییر داده solarization را به مجموعه data transformation ها، با احتمال ۰.۱ اضافه کنید و پس از آن، همه data augmentation ها را خاموش و تنها crop را نگاه دارید و مجدداً مدل را آموزش دهید. تفاوت نتیجه ها را گزارش کنید.

بخش سوم

در SimCLR هر عکس باید تغییر یافته دیگر خود را بیابد و به آن نزدیک شود و از بقیه دور شود، به علاوه، با استفاده از masking با خودش مقایسه نمی شود. حال در این مرحله، قصد داریم سخت ترین (نزدیک ترین) نمونه غیر یکسان با عکس را هم از تابع هزینه حذف کنیم (با masking). یعنی برای مثال اگر در یک batch با اندازه ۳ بخواهم توضیح دهم، اینطور می شود:

Given $Img_{1,2,3}$, we will augment them and get $x_1, x_2, y_1, y_2, z_1, z_2$ representations:

Vanilla loss: $d(x_1, x_2) / (d(x_1, x_2) + d(x_1, y_1) + d(x_1, y_2) + d(x_1, z_1) + d(x_1, z_2))$

Remove-Hardest loss: let's assume $d(x_1, y_2) < d(x_1, z_1) < d(x_1, z_2) < d(x_1, y_1)$, so y_2 is the closest to x_1 and we will remove y_1 and y_2 from the denominator:

Loss = $d(x_1, x_2) / (d(x_1, x_2) + d(x_1, z_1) + d(x_1, z_2))$

نتیجه پیاده سازی و اجرا را گزارش کنید.

سوال سوم – Contrastive Predictive Coding

در سوال سوم در ابتدا مقاله زیر را مطالعه کنید تا اطلاعات کافی راجع به روش " Moment Contrast" به دست آورید.

Representation Learning with Contrastive Predictive Coding

در این روش برای buffer کردن نمونه های negative از یک feature representation استفاده می شود که ساختاری مانند feature representation اصلی دارد با این تفاوت که بروز رسانی پارامتر ها آن متفاوت انجام می شود.

در نتیجه می توانیم تعداد نمونه های negative را بیشتر کنیم و در InfoNCE loss وارد کنیم که طبیعتاً در نهایت دقت بالاتری خواهیم داشت. با استفاده از این روش و همچنین با اتکا بر ساختار

- Resnet18
- VGG16

با استفاده از تعداد صد ایاک، feature representation مناسب بر روی دادگان CIFAR10 را

با اعمال

- Rotation
- Exemplar

آموزش دهید. و همچنین برای طبقه بندی کردن در نهایت میتوانید از linear classifier استفاده کنید.

در نهایت شما باید دقت مدل را به ازای ساختار های مختلف و روش های augmentation مختلف گزارش کنید.

سوال چهارم – Downstream task

بخش اول

در آخرین مسئله قصد داریم که downstream task را با دو روش linear classifier و fine tuning انجام دهیم. به این منظور با طراحی یک معماری مناسب (مثلا autoencoder) به عنوان pretext، مجموعه دادگان cifar10 را با اتکا بر روش های augmentation زیر به صورت self-supervised آموزش دهید.

- Adding Gaussian noise
- JIGSAW
- Gray scale image to RGB

سپس Feature representation مناسب را که نتیجه self-supervision میباشد، پیدا کنید . سپس برای evaluation کردن از روش های زیر استفاده کنید:

- Linear classifier
- Fine-tuned models:
 - resnet18
 - VGG16

سپس مقایسه همه روش های اعمال شده را در جدول کنار هم گزارش دهید.

بخش دوم

حالا برای قسمت pretext در سوال بالا، از یک معماری auto-regressive generative model مانند PixelCNN استفاده کنید. برای مثال می توانید representation مناسب را در این تعریف کنید که نیمه بالایی representation، نیمه پایینی آن را پیش بینی کند. سپس تعدادی از خروجی های ساخته شده را نمایش دهید و برای evaluate کردن مدل به صورت دلخواه از linear classifier یا fine-tuned model استفاده کنید.

نکات:

- مهلت تحویل این پروژه ۳۱ خرداد ۱۴۰۰ 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 فرضیهایی که برای پیاده سازی‌ها و محاسبات خود در نظر می‌گیرید را در گزارش ذکر کنید.
- در گزارش خود برای تصاویر زیرنویس و برای جداول هم بالانویس اضافه کنید.
- در متن گزارش الزامی به ارائه توضیح جزئیات خط به خط کد در گزارش نیست. اما باید نتایج بدست آمده را ارائه دهید و تحلیل کنید.
- 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.
- برای انجام پروژه‌ها، تنها زبان برنامه نویسی مجاز Python است.
- در صورت مشاهده تقلب، امتیاز تمامی افراد شرکت کننده در آن ۱۰۰- لحاظ می‌شود.
- نحوه محاسبه تاخیر به این شکل است: پس از اتمام مهلت اصلی ارسال، به مدت سه روز (تا ۳ تیر) بارگذاری با تاخیر ممکن است، اما به ازای هر روز ۲۰ درصد از نمره کسر خواهد شد؛ در نهایت و پس از بازه تاخیر، ارسال ممکن نیست و نمره تکلیف صفر خواهد شد.
- لطفاً گزارش، فایل کدها و سایر ضمایم مورد نیاز را با فرمت زیر در سامانه مدیریت دروس بارگذاری نمایید.

PROJECT#_[Lastname]_[StudentNumber].zip

- در صورت وجود ابهام و پرسش، می‌توانید از طریق رایانامه‌های زیر با دستیاران آموزشی مربوطه آقایان سپهر سامنی (سوال‌های اول و دوم) و حمیدرضا هاشم‌پور (سوال سوم و چهارم) در تماس باشید:

hamidreza.hashemp@ut.ac.ir

sepehr.sameni@gmail.com