

شبکه‌های عصبی
دانشگاه فردوسی مشهد
تمرین ششم

نیمسال دوم تحصیلی ۱۴۰۲-۱۴۰۳

مهلت ارسال: ۱۴۰۳/۰۴/۲۰

گروه مهندسی کامپیوتر

در این تمرین قصد داریم چند روش مختلف پیش‌تعلیم^۱ بدون برچسب را برای دیتاست Cifar10 مقایسه کنیم. توضیحات در مورد مدل‌های مختلف و Masked AutoEncoders در ویدئو تکمیلی برای این تمرین در سامانه مجازی آموزشی قرار داده شده، پیش از مطالعه موارد خواسته شده برای پیاده‌سازی این ویدئو را نگاه کنید. همچنین قطعه کدی به شما داده شده است که برای انجام تمرین آن را تکمیل کنید.

۱) آزمایشات

در هرکدام از آزمایشات، نمودار تابع زیان برای داده‌های آموزشی و داده‌های تست را در طول آموزش رسم کنید و گزارش دهید. برای قسمت ۲.۱ و ۴.۱ نمودار دقت برای داده‌های آموزشی و داده‌های تست را در طول آموزش رسم کنید و گزارش دهید. برای قسمت ۱.۱ چندین تصویر از بخش Test دیتاست به همراه تصویر بازسازی شده توسط مدل برای آن تصویر را رسم کنید و گزارش دهید. (این تصاویر را برای ایپاک‌های مختلف یادگیری گزارش دهید.) همچنین برای قسمت ۳.۱ چندین تصویر، Test، به همراه تصویر معادل Mask شده که به مدل داده می‌شود، به همراه تصویر بازسازی شده توسط مدل را برای ایپاک‌های مختلف رسم کنید و گزارش دهید. (مانند شکل ۱) برای پیاده‌سازی از قطعه کد داده شده استفاده کنید و برای ماجول‌های Encoder مربوط به Transformer از ماجول‌های از پیش آماده شده pytorch استفاده کنید (توضیحات بیشتر در ویدئو).

۱.۱) آموزش AutoEncoder معمولی

در اولین بخش یک AutoEncoder معمولی برای بازسازی تصاویر ورودی شبکه آموزش دهید. از ماجول Encoder مربوط به Transformer توضیح داده شده در ویدئو برای این AutoEncoder استفاده کنید. مدل را با داده‌های Train از دیتاست آموزش دهید و با داده‌های Test کیفیت تصاویر بازسازی شده توسط مدل را بررسی کنید.

۲.۱) طبقه‌بندی با کمک Encoder

از Encoder مدل آموزش داده شده در قسمت قبلی برای طبقه‌بندی تصاویر استفاده کنید. به این صورت که یک لایه Fully Connected به خروجی مربوط به cls_token از Encoder از پیش تعلیم داده شده بچسبانید (توضیحات بیشتر در ویدئو تکمیلی) و با داده‌های Train از دیتاست مدل را برای آموزش مدل با تسک Classification استفاده کنید. دقت شود هم می‌توانید Encoder از پیش تعلیم شده را که به عنوان Extractor Feature در مدل جدید کار می‌کند را در این قسمت آموزش دهید و هم می‌توانید پارامترهای آن را Freeze کنید.

^۱Pre-training

۳.۱) Masked AutoEncoder

مدل AutoEncoder را به طور Masked شده برای بازسازی patch های mask شده در تصویر ورودی آموزش دهید. توضیحات بیشتر برای مدل Masked AutoEncoder در ویدئو تکمیلی داده شده است. مدل را با داده‌های Train از دیتاست آموزش دهید و با داده‌های Test کیفیت تصاویر بازسازی شده توسط مدل را بررسی کنید.

۴.۱) طبقه بندی با کمک Masked AutoEncoder

مشابه کاری که باید در قسمت ۲.۱ برای مدل ۱.۱ انجام شود، از Encoder آموزش داده شده در مدل قسمت ۳.۱ برای طبقه بندی تصاویر استفاده کنید. به این صورت که یک لایه Fully Connected به خروجی مربوط به cls_token از Encoder از پیش تعلیم داده شده بچسبانید (توضیحات بیشتر در ویدئو تکمیلی) و با داده‌های Train از دیتاست مدل را برای آموزش مدل با تسک Classification استفاده کنید. دقت شود هم می‌توانید Encoder از پیش تعلیم شده را که به عنوان Extractor Feature در مدل جدید کار می‌کند را در این قسمت آموزش دهید و هم می‌توانید پارامترهای آن را Freeze کنید.



شکل ۱: از هر سه تصویر در یک ردیف، تصویر سمت چپ تصویر mask شده که به مدل داده می‌شود است. تصویر میانی بازسازی شده توسط مدل است و تصویر سمت راست تصویر اصلی است. تصاویر بازسازی شده برای مدل Masked AutoEncoder که در ویدئو توضیح داده شده است.