فاز اول | بینایی ماشین

چالش یادگیری ماشین

Machine Learning Challenge

برگزارکننده:

دستیاران آموزشی درس هوش مصنوعی دانشگاه صنعتی شریف

بهار ۱۴۰۲




مقدمه

در این چالش شما با آموزش مدل‌های دسته‌بندی تصاویر و پس از آن، تنومند کردن آن‌ها در برابر نمونه‌های خصمانه دست و پنجه نرم خواهید کرد. چالش به صورت کلی از ۳ بخش تقسیم شده است که به توضیح هر کدام از آن‌ها خواهیم پرداخت.

۱. آموزش دسته‌بند

یک مجموعه داده از تصاویر ۸ حیوان مختلف برای این چالش تهیه شده است. داده های آموزش را می توانید از این لینک دریافت کنید. هدف دسته بندی تصاویر به یکی از ۸ کلاس متناظر با حیوانی که آن تصویر نشان می دهد است. در این مجموعه داده تعداد زیادی تصویر با برچسب مربوطه وجود دارد که قسمت آموزش آن به شرکت کنندگان داده می شود. در این بخش باید با استفاده از شبکه‌های عصبی همچون CNNها، دسته‌بندی پیاده کنید تا برچسب صحیح مربوط به تصویر ورودی را تشخیص دهد.

۲. تنومندسازی دسته‌بند

دسته‌بندها در حالت معمول روی ورودی‌های تغییر یافته خوب عمل نمی‌کنند. در این مورد می‌توانید در انتهای همین داک بخش خصمانه مطالعه کنید. شما باید مدل خود را به نحوی آموزش دهید تا بتواند در برابر نمونه‌های خصمانه مقاومت داشته باشد.

جزئیات آموزش خصمانه: مقدار نرم بی‌نهایت برای نویز اضافه‌شده به تصویر می‌تواند حداکثر 8/255 باشد (برای مقادیر بین ۰ تا ۲۵۵، بنابراین اگر مقیاس عددی پیکسل‌ها را تغییر داده‌اید باید این محدودیت را به همان بازه مقیاس کنید). در حمله PGD مقدار مجاز برای حلقه داخلی حداکثر ۳ می‌باشد.

۳. حمله به سایر مدل‌ها

در این بخش باید سعی کنید نمونه‌های خصمانه بسازید تا به مدل‌های رقبای خود حمله کنید. برای این منظور، شما با فرض اینکه مدل حریف را دارید، یک حمله جعبه سفید طراحی می‌کنید تا ورودی‌ها را گرفته و نمونه‌های خصمانه تولید کند.

کد شما در سرور مسابقه اجرا شده و نمونه‌های خصمانه برای مدل‌های رقیبان ایجاد شده و دقت مدلشان محاسبه می‌شود.

فاز اول: دسته‌بندی تصاویر

در فاز اول هدف طراحی و پیاده‌سازی یک مدل برای دسته‌بندی است. مدل طراحی‌شده و آموزش داده شده توسط هر شرکت‌کننده پس از ارسال، روی داده‌های تست توسط معیار دقت (accuracy) ارزیابی می‌شود که معنای دقیق آن نسبت تعداد پاسخ‌های درست مدل به تعداد کل داده‌هاست. شرکت‌کنندگان بر اساس دقت مدلشان رتبه‌بندی می‌شوند و در انتهای فاز اول مسابقه، **۸ گروه** شرکت‌کننده با امتیاز برتر انتخاب و به فاز دوم راه پیدا می‌کنند. دقت کنید که امتیازات و رتبه‌بندی مسابقه تنها با ۳۰ درصد از داده‌های تست اعلام می‌شود و پس از اتمام زمان این فاز، شرکت‌کنندگان بر اساس دقت مدلشان روی ۷۰ درصد دیگر داده‌های تست رتبه‌بندی خواهند شد.

نحوه ارسال

برای ارسال مدل خود باید یک کد پایتون به فرمت زیر با نام `model.py` داشته باشید:

```
class ClassificationModel(nn.Module):
    def __init__(self):
        # code for model initialization

    # Image batch should be a 4D tensor of size N * C * H * W
    # returns a N * n_classes tensor
    def forward(self, image_batch):
        # computation of the model's output
        return prediction
```

ورودی مدل باید یک تنسور ۴ بعدی که به صورت

`تعداد عکس‌ها * تعداد کانال‌های ورودی (۳) * طول عکس * عرض عکس`

باشد. خروجی مدل یک تنسور به ابعاد

`تعداد عکس‌ها * تعداد کلاس‌های دسته‌بندی`

می‌باشد که امتیاز مدل برای هر کلاس را برای هر عکس ورودی نشان می‌دهد.

همچنین یک فایل `checkpoint.pth` نیاز هست که وزن‌های شبکه آموزش داده شده باید داخل آن قرار گیرد.

دقت کنید مدل باید به صورت زیر قابل خواندن و استفاده باشد:

```
model = ClassificationModel()
model.load_state_dict(torch.load("checkpoint.pth"))
```

در نهایت یک فایل `requirements.txt` باید وجود داشته باشد که در آن لیست کتابخانه‌های مورد نیاز برای اجرای کد شما وجود داشته باشد. این **سه فایل** را در کنار هم در یک فایل زیپ ارسال کنید.

محدودیت‌های زمان اجرا

حداکثر زمان و حافظه قابل قبول برای اجرای روی یک batch داده به اندازه ۱۶ با تصاویر به سایز `۲۲۴ * ۲۲۴`:

```
GPU Memory: 4GB
RAM: 4GB
Time: 0.5s
```

محدودیت‌های پیاده‌سازی

- برای پیاده‌سازی مدل حتما از کتابخانه PyTorch استفاده کنید.
- استفاده از مدل‌های آماده در PyTorch برای پیاده‌سازی دسته‌بند مانعی ندارد.