

شبکه‌های مولد تقابلی

سوال ۱

در این تمرین، هدف بررسی معماری‌ها و قابلیت‌های مختلف شبکه‌های مولد تقابلی (GAN) در تولید تصاویر واقعی و با کیفیت است. در این پروژه، شما وظیفه دارید یک مدل پایه GAN را پیاده‌سازی کرده، سپس با اعمال تغییرات ساختاری و تکنیکی، عملکرد آن را بهبود دهید و نتایج را تحلیل نمایید.

مجموعه‌داده:

از مجموعه‌داده CIFAR-10 استفاده نمایید. این مجموعه شامل ۶۰۰۰۰ تصویر رنگی 32×32 در ۱۰ کلاس مختلف است که در این پروژه صرفاً برای یادگیری مولد (بدون برچسب) به کار می‌رود.

۱ - مدل پایه‌ی DCGAN را مطابق معماری معرفی شده در مقاله اصلی پیاده‌سازی نمایید. پس از آموزش مدل برای چندین تکرار، تصاویر تولیدشده توسط مولد را ذخیره کرده و با تصاویر واقعی مقایسه نمایید.

۲ - معیارهای کیفی و کمی زیر را برای ارزیابی مدل استفاده نمایید:

Inception Score (IS) –

Frechet Inception Distance (FID) –

– مقایسه دیداری با تصاویر واقعی

۳ - آزمایش‌های حساسیت به پارامترهای مختلف را انجام داده و تأثیر آن‌ها را بررسی و تحلیل نمایید:

– اندازه بردار تویز ورودی

– استفاده از Batch Normalization در مولد یا متمایز‌کننده

– استفاده از Label Smoothing

– انتخاب توابع فعال‌سازی متفاوت (ReLU, LeakyReLU, Tanh)

۴ - مدل (cGAN) را پیاده‌سازی نمایید که قابلیت کنترل کلاس تصویر خروجی با استفاده از برچسب را داشته باشد. برای این منظور از برچسب‌های CIFAR-10 استفاده نمایید و تصاویری از چند کلاس مشخص مانند «گربه»، «ماشین» و «پرنده» تولید کنید.

۵ - مدل (WGAN-GP) را پياده‌سازی نمایيد. تفاوت Wasserstein GAN with Gradient Penalty (WGAN-GP) در تولید تصاویر، پايداري آموزش و معيارهای عددی تحليل نمایيد.

۶ - با استفاده از t-SNE یا UMAP، توزيع تصاویر واقعی و تولیدشده در فضای ويژگی را رسم نموده و بررسی کنيد که آيا توزيعها به هم نزدیک هستند یا خير.

۷ - تحليل مفهومی انجام دهيد:

- نقاط قوت و ضعف GAN‌ها در مقایسه با ديگر مدل‌های مولد مانند VAEs

- چالش‌های آموزش GAN (نظير ناپايداري، فروپاشی حالت، دشواری همگرائي)

- راهکارهای پيشنهادی برای بهبود فرآيند يادگيري و تثبيت آموزش

سوال ۲

این سوال، امتيازی است.

در اين بخش قصد داريم با استفاده از ساختار معرفی شده در مقاله GANomaly^۱، يك مدل برای تشخيص ناهنجاري در تصاویر پياده‌سازی کنيم. ساختار کلی اين مدل مبتنی بر معماری خودکدگذار است. هدف اين مدل بازسازی تصاویر ورودی و اندازه‌گيري ميزان ناهنجاري بر اساس تفاوت ميان تصوير ورودی و بازسازی شده، بازنمياني‌هاي فشرده آن‌ها و همچنین خروجی تفكيك‌کننده است.

۱ - با استفاده از کتابخانه torchvision مجموعه داده‌های MNIST را دریافت کرده و تعدادی از تصاویر آن را نمایش دهيد.

۲ - يك خودکدگذار طراحی کنيد که شامل کدگذار و کدگشای قرينه باشد. می‌توانيد از ساختار مقاله (DCGAN) يا ساختار دلخواه خود استفاده کنيد. ساختار نهايی باید قابلیت بازسازی مناسب تصوير را داشته باشد. معماری و کد آن را در گزارش خود شرح دهيد.

۳ - تابع L_{con} را پياده‌سازی نمایيد که ميزان اختلاف ميان تصوير اصلی و بازسازی شده را با استفاده از معيارهایي مانند MAE یا MSE اندازه‌گيري کند.

۴ - کدگذار طراحی شده را به صورت اشتراكی برای سه بخش استفاده کنيد:

^۱مقاله مذكور به فايل تعريف پروژه ضميمه شده است.

^۲(Auto-encoder)

- ۱ - بازنمایی فشرده تصویر ورودی (در مرحله بازسازی)
- ۲ - بازنمایی فشرده تصویر بازسازی شده (به عنوان کدگذار دوم)
- ۳ - استفاده از خروجی کدگذار به عنوان ورودی به بخش تفکیک کننده
- ۴ - توابع خطای L_{adv} و L_{enc} را تعریف کرده و کاربرد آنها را توضیح دهید. سپس با ترکیب این توابع L_{con} ، تابع هدف نهایی مدل را بر اساس مقاله تعریف کنید.
- ۵ - برای محاسبه امتیاز ناهنجاری، می‌توانید از ترکیب توابع خطای L_{adv} یا L_{enc} تعریف شده در مقاله استفاده کنید. روش انتخابی خود را توضیح دهید.
- ۶ - فرآیند آموزش را برای تشخیص ناهنجاری در کلاس‌های مختلف مجموعه MNIST پیاده‌سازی کنید:

 - برای هر کلاس (مثلاً عدد ۰) تنها نمونه‌های آن کلاس را برای آموزش استفاده کنید.
 - در زمان تست، نمونه‌هایی از همه کلاس‌ها را ورودی دهید و بررسی کنید که آیا مدل تنها آن کلاس را به عنوان هنجار تشخیص می‌دهد یا خیر.

- ۷ - برای هر کلاس، مقدار AUC-ROC را محاسبه کرده و در گزارش ارائه دهید. همچنین نمونه‌هایی از تصاویر ورودی و تصاویر بازسازی شده (چه در موارد موفق و چه در موارد ناموفق) را نمایش دهید و تحلیل مختصری از آنها ارائه نمایید.

در انجام تمرینات به نکات زیر توجه فرمایید.

- ۱ - پیاده‌سازی با زبان Python و چارچوب کاری PyTorch انجام شود. استفاده از torchvision برای بارگذاری داده‌ها مجاز است.
- ۲ - گزارش نهایی باید شامل تحلیل دقیق عملکرد، نمودارهای Loss، نتایج عددی FID/IS تصاویر تولیدشده و مقایسه با نمونه‌های واقعی باشد.
- ۳ - در صورت استفاده از مدل‌های از پیش آماده یا معماری‌های مقاله‌ای، منابع را به‌طور کامل ذکر کنید.
- ۴ - ارسال فایل فشرده شامل کد و گزارش با نام StudentID_HW07.zip در موعد مقرر الزامی است. در غیر این صورت نمره تمرین لحاظ نخواهد شد.
- ۵ - کپی‌برداری از دیگر دانشجویان یا استفاده مستقیم از کدهای آماده بدون درک کامل، موجب حذف کل نمره تمرین خواهد شد.