

شبكههاى مؤلء آءابلى

سؤال ۱

ءر اىن آمرىن؁ هءف بررسى معمارىها و قابلىءهاى مؤلف شبكههاى مؤلء آءابلى (GAN) ءر ءولء ءصاوءر واقعى و با كىفىء اسء. ءر اىن ٱروژه؁ شما وظىفه ءارءء يك مءل ٱاه GAN را ٱىاءهسازى كءه؁ سٱس با اعمال آءىىراء ساختارى و آكنىكى؁ عملكرء آن را بهبوء ءهءء و نءاءى را آءلىل نماىءء.

مجموعه ءاءه:

از مجموعه ءاءه CIFAR-10 اسءفءه نماىءء. اىن مجموعه شامل ۶۰۰۰۰ ءصوءر رنگى ۳۲×۳۲ ءر ۱۰ كلاس مؤلف اسء كه ءر اىن ٱروژه صرفاً برائى ياءءءبرى مؤلء (بءون برآسب) به كار مىرءء.

۱ - مءل ٱاهى DCGAN را مطابق معمارى معرفى شءه ءر مقالة اصلى ٱىاءهسازى نماىءء. ٱس از آموزش مءل برائى آءءءن آءكار؁ ءصاوءر ءولءءشءه ءوسط مؤلء را ءآىره كءه و با ءصاوءر واقعى مقابسه نماىءء.

۲ - معيارهاى كىفى و كمى زىر را برائى ارزىابى مءل اسءفءه نماىءء:

– Inception Score (IS)

– Frechet Inception Distance (FID)

– مقابسه ءىءارى با ءصاوءر واقعى

۳ - آزمابشهاى حساسىء به ٱارامءرهاى مؤلف را انآام ءاءه و آأىءر آنها را بررسى و آءلىل نماىءء:

– انءازه برءار نوءز وروءى

– اسءفءه از Batch Normalization ءر مؤلء يا مءماىزكنءه

– اسءفءه از Label Smoothing

– انءآاب ءوابع فعالسازى مءفاوء (ReLU, LeakyReLU, Tanh)

۴ - مءل Conditional GAN (cGAN) را ٱىاءهسازى نماىءء كه قابلىء كنءرل كلاس ءصوءر آروءى با اسءفءه از برآسب را ءاشءه باشء. برائى اىن منظور از برآسبهاى CIFAR-10 اسءفءه نماىءء و ءصاوءرى از آءن كلاس مشآص مانءء «ءر به»؁ «ماشىن» و «ٱرءه» ءولءء كنءء.

- ۵ - مدل Wasserstein GAN with Gradient Penalty (WGAN-GP) را پیاده‌سازی نمایید. تفاوت عملکرد این مدل را با DCGAN در تولید تصاویر، پایداری آموزش و معیارهای عددی تحلیل نمایید.
- ۶ - با استفاده از t-SNE یا UMAP، توزیع تصاویر واقعی و تولیدشده در فضای ویژگی را رسم نموده و بررسی کنید که آیا توزیع‌ها به هم نزدیک هستند یا خیر.
- ۷ - تحلیل مفهومی انجام دهید:

- نقاط قوت و ضعف GANها در مقایسه با دیگر مدل‌های مولد مانند VAEs
- چالش‌های آموزش GAN (نظیر ناپایداری، فروپاشی حالت، دشواری همگرایی)
- راهکارهای پیشنهادی برای بهبود فرآیند یادگیری و تثبیت آموزش

سوال ۲

این سوال، امتیازی است.

در این بخش قصد داریم با استفاده از ساختار معرفی شده در مقاله GANomaly^۱، یک مدل برای تشخیص ناهنجاری در تصاویر پیاده‌سازی کنیم. ساختار کلی این مدل مبتنی بر معماری خودکدگذار^۲ است. هدف این مدل بازسازی تصاویر ورودی و اندازه‌گیری میزان ناهنجاری بر اساس تفاوت میان تصویر ورودی و بازسازی شده، بازنمایی‌های فشرده آن‌ها و همچنین خروجی تفکیک‌کننده است.

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

- ۲ - یک خودکدگذار طراحی کنید که شامل کدگذار و کدگشای قرینه باشد. می‌توانید از ساختار مقاله (DCGAN) یا ساختار دلخواه خود استفاده کنید. ساختار نهایی باید قابلیت بازسازی مناسب تصویر را داشته باشد. معماری و کد آن را در گزارش خود شرح دهید.

- ۳ - تابع L_{con} را پیاده‌سازی نمایید که میزان اختلاف میان تصویر اصلی و بازسازی شده را با استفاده از معیارهایی مانند MSE یا MAE اندازه‌گیری کند.

- ۴ - کدگذار طراحی شده را به صورت اشتراکی برای سه بخش استفاده کنید:

^۱مقاله مذکور به فایل تعریف پروژه ضمیمه 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
- ۵ - توابع خطای L_{adv} و L_{enc} را تعریف کرده و کاربرد آن ها را توضیح دهید. سپس با ترکیب این توابع با L_{con} ، تابع هدف نهایی مدل را بر اساس مقاله تعریف کنید.
- ۶ - برای محاسبه امتیاز ناهنجاری، می توانید از ترکیب توابع خطا یا تابع تعریف شده در مقاله استفاده کنید. روش انتخابی خود را توضیح دهید.
- ۷ - فرآیند آموزش را برای تشخیص ناهنجاری در کلاس های مختلف مجموعه MNIST پیاده سازی کنید:
 - برای هر کلاس (مثلاً عدد ۰) تنها نمونه های آن کلاس را برای آموزش استفاده کنید.
 - در زمان تست، نمونه هایی از همه کلاس ها را ورودی دهید و بررسی کنید که آیا مدل تنها آن کلاس را به عنوان هنجار تشخیص می دهد یا خیر.
- ۸ - برای هر کلاس، مقدار AUC-ROC را محاسبه کرده و در گزارش ارائه دهید. همچنین نمونه هایی از تصاویر ورودی و تصاویر بازسازی شده (چه در موارد موفق و چه در موارد ناموفق) را نمایش دهید و تحلیل مختصری از آن ها ارائه نمایید.

در انجام تمرینات به نکات زیر توجه فرمایید.

- ۱ - پیاده سازی با زبان Python و چارچوب کاری PyTorch انجام شود. استفاده از torchvision برای بارگذاری داده ها مجاز است.
- ۲ - گزارش نهایی باید شامل تحلیل دقیق عملکرد، نمودارهای Loss، نتایج عددی FID/IS، تصاویر تولید شده و مقایسه با نمونه های واقعی باشد.
- ۳ - در صورت استفاده از مدل های از پیش آماده یا معماری های مقاله ای، منابع را به طور 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 - کپی برداری از دیگر دانشجویان یا استفاده مستقیم از کدهای آماده بدون درک کامل، موجب حذف کل نمره تمرین خواهد شد.