

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



شبکه های عصبی و یادگیری عمیق

مینی پروژه شماره ۳

دی ماه ۹۸

سوال ۱

به سوالات زیر در مورد شبکه‌های GAN پاسخ دهید.

I- در میان مدل‌های ارائه شده برای GAN یکی از جدیدترین مدل‌ها Self-Attention Generative Adversarial Network است که بهترین نتیجه را در تولید تصاویر برای مجموعه داده‌های چند کلاسه داشته است. مقاله مورد نظر را در [اینجا](#) می‌توانید پیدا کنید. انگیزه استفاده از self-attention در کار با تصاویر چیست؟ نویسندگان این مقاله چگونه از self-attention در مدل خود استفاده کرده‌اند؟

II- تفاوت‌های Autoencoders و Variational Autoencoders را بیان کرده و مراحل پیاده‌سازی یک VAE را توضیح دهید.

سوال ۲ – DCGAN

یکی از مشهورترین ساختارهای GAN، معماری DCGAN است که در طراحی این معماری از لایه‌های CNN در ساختار مولد و تفکیک‌کننده استفاده شده است. لطفاً به منظور آشنایی بیشتر با این معماری ابتدا [مقاله مربوطه](#) را مطالعه کرده، سپس به سوالات زیر پاسخ دهید.

I- ساز و کار طراحی معماری DCGAN را با جزییات توضیح دهید و دلیل استفاده از موارد زیر در طراحی ساختار DCGAN را بیان کنید:

(a) استفاده از Strided Convolution و Fractional Strided Convolution برای Upsampling و Downsampling

(b) استفاده از Batch Normalization

(c) استفاده از تابع فعال‌ساز Leaky Relu

(d) استفاده از روش بهینه‌سازی ADAM

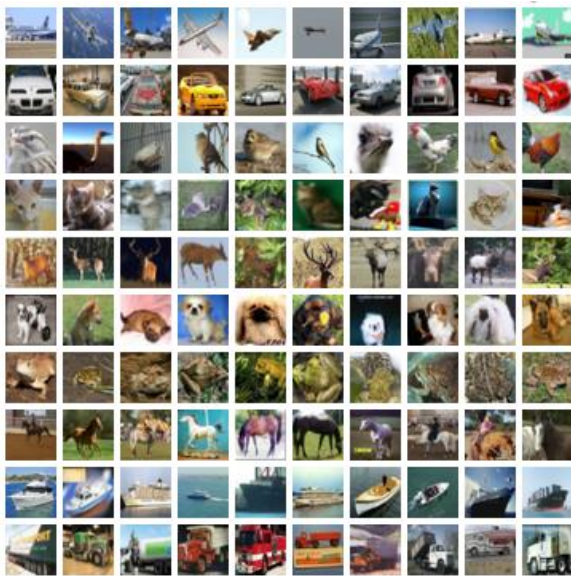
II- برنامه‌ای بنویسید تا با استفاده از معماری DCGAN توانایی تولید عکس از کلاس‌های مختلف برای مجموعه داده‌های [CIFAR-10](#) و [Fashion MNIST](#) را داشته باشد. برای هر مجموعه از داده‌ها موارد زیر را گزارش دهید:

(a) توپولوژی لایه‌های مولد و تفکیک‌کننده

(b) نحوه ایجاد نویز

(c) نمودارهای Loss و Accuracy برای هر دو بخش مولد و تفکیک کننده
 (d) نمونه نتایج شبکه مولد به ازای epochهای مضرب 10 (تعداد epoch را برابر با ۱۰۰ قرار دهید)

توجه کنید که تصاویر تولیدی باید به صورت یک ماتریس ۱۰ در ۱۰ به صورت زیر باشند که بدین منظور میتوانید از کتابخانه [Pillow](#) استفاده نمایید:



CIFAR-10



Fashion-MNIST

سوال ۳ – Ways to improve GAN performance

شبکه‌های Gan عادی دارای معایب متعددی از جمله مشکل ناپایداری شبکه، مشکل Vanishing Gradient هنگام استفاده از تابع هزینه MinMax و Mode collapse می‌باشند. در ادامه تکنیک‌هایی جهت رفع مشکلات مذکور ارائه خواهد شد.

I- جهت پایدارسازی شبکه‌های GAN میتوان از دو تکنیک ساده [One-sided Label Smoothing](#) و [اضافه کردن نویز](#) استفاده نمود. لطفا مقالات مربوط به هر تکنیک را مطالعه کرده و نحوه پیاده‌سازی هریک را با جزییات بیان کنید، سپس با اصلاح شبکه DCGAN طراحی شده در سوال قبل، موارد خواسته شده در قسمت II سوال دوم را تنها برای مجموعه داده CIFAR-10 ارائه دهید.

II- در GAN های معمولی، یکی از اقداماتی که در جهت بهبود مشکل Vanishing gradient انجام می شود، استفاده از Wasserstein loss می باشد این تابع هزینه برای اولین بار در شبکه های [WGAN](#) مطرح شد. با مطالعه مقاله مربوطه ابتدا توضیح دهید Wasserstein loss چیست و تفاوت آن با loss مورد استفاده در GAN های معمولی چه می باشد، سپس ساختار WGAN را پیاده سازی کرده و موارد خواسته شده در قسمت II سوال دوم را تنها برای مجموعه داده CIFAR-10 ارائه دهید.

III- همانطور که گفته شد، یکی دیگر از مشکلات شبکه های GAN، بحث Mode collapse میباشد. یکی از روش های رایج جهت رفع این مشکل، استفاده از لیبل های هر کلاس در پروسه یادگیری است. ساختار [ACGAN](#) نیز از همین روش جهت بهبود کیفیت تصاویر تولیدی استفاده می کند. ابتدا مقاله مربوط به ساختار ACGAN را مطالعه کنید و سازوکار آن را شرح دهید، سپس این ساختار را پیاده سازی کرده و موارد خواسته شده در قسمت II سوال دوم را تنها برای مجموعه داده CIFAR-10 ارائه دهید. در این بخش علاوه بر ارائه عکس های خروجی شبکه، لیبل تخمین زده شده مربوط به هر عکس را نیز ارائه دهید.

IV- نتایج بخش های I، II و III را با سوال دوم مقایسه کنید.

نکات مهم پیاده‌سازی:

- استفاده از هر یک از framework های Tensorflow، Keras و Pytorch مجاز می‌باشد اما به دلیل سهولت در پیاده‌سازی، استفاده از Keras به شدت توصیه می‌گردد.
- به دلیل زمان زیاد اجرای هرکد، اکیدا توصیه می‌شود از Google Colab استفاده کنید و یا با استفاده از کتابخانه CUDA کدها را بر روی GPU اجرا کنید.

دیگر نکات:

- مهلت تحویل قانونی! این تمرین ۲۰ دی می‌باشد اما ارسال تمرین بدون در نظر گرفتن جریمه تا ۴ بهمن نیز امکان‌پذیر می‌باشد (تاریخ مذکور به هیچ عنوان قابل تمدید نخواهد بود).
- گزارش شما در فرآیند تصحیح از اهمیت ویژه‌ای برخوردار است. لطفا تمامی نکات و فرض‌هایی که برای پیاده‌سازی‌ها و محاسبات خود در نظر می‌گیرید را در گزارش ذکر کنید.
- در صورت مشاهده‌ی تقلب نمرات تمامی افراد شرکت کننده در آن صفر لحاظ می‌شود.
- استفاده از کدهای آماده برای تمرین‌ها مجاز نمی‌باشد. برای مینی پروژه‌ها فقط برای قسمت‌هایی از کد و به عنوان راهنمایی برای پیاده‌سازی می‌توانید از کدهای آماده استفاده کنید.
- در صورت وجود هرگونه ابهام یا مشکل می‌توانید از طریق رایانامه‌ی زیر با دستیار آموزشی مربوطه در تماس باشید:

حمیدرضا آفتابی hamid.aftabi@gmail.com

دانیال صدریان زاده daniyal.sadriyan@ut.ac.ir