امیرمحمد عزتی 4092212269 - تمرین سری ششم تصاویر پزشکی - تئوری

## سوال ①

۱. از لحاظ معماری:

Diffusion model: مدل دیفیوژن از لحاظ معماری از دو بخش forward process و reverse process تشکیل شده است که بخش reverse آن می‌تواند از یک نویز گاوسی به یک عکس جدید برسد.

GAN: GAN از دو شبکه generator و discriminator تشکیل شده است. شبکه discriminator باید داده‌های واقعی را از داده‌های فیک تشخیص دهد و شبکه generator باید از یک توزیع p(z) عکس‌هایی تولید کند که شبکه discriminator را فریب دهد.

VAE: اتوانکودر variational از دو بخش encoder و decoder تشکیل شده است. این شبکه طوری کلی یک ورودی عکس را به خودش در خروجی تبدیل می‌کند. امّا در فضای latent خود میانگین و واریانس توزیع داده‌ها را به ما می‌دهد که می‌توانیم از آن نمونه بگیریم و با بخش decoder عکس جدید تولید کنیم.

روش آموزش:

هر سه مدل بصورت Supervised آموزش داده می‌شوند.

Diffusion: دیفیوژن بتدریج به ورودی نویز اضافه می‌کند تا مرحله‌ای که به یک نویز خالص برسد و در فرایند reverse، نویز را به عکس تبدیل می‌کند.

GAN: ابتدا discriminator روی داده‌های واقعی با لیبل 1 آموزش داده می‌شود سپس generator با استفاده از ورودی با توزیع p(z) سعی می‌کند عکس تولید کند و شبکه discriminator روی داده‌های خروجی generator (که فیک هستند) با لیبل 0 آموزش داده می‌شود و خروجی 0 از discriminator به generator داده می‌شود و باعث می‌شود generator عکس‌هایی تولید کند که شبکه discriminator را فریب دهد.

VAE : در این شبکه ، ورودی و خروجی هردو یک عکس داده می‌شود و این شبکه باید یاد بگیرد عکس را تبدیل به feature map کرده و دوباره از feature map ، reconstruct کند.

کیفیت و تنوع نمونه‌ها :

GAN می‌تواند عکس‌های باکیفیت بالا و واقعی تولید کند و همچنین عکس‌های متنوع ارائه کند. امّا ممکن است دچار mode collapse شود و چند حالت از توزیع داده را تولید کند و بقیه را نادیده بگیرد.

VAE می‌تواند نمونه‌های متنوع تولید کند امّا خروجی‌هایش ممکن است نویزی و blurry باشد و نمی‌تواند داده‌های پیچیده با فضای latent ابعاد بالا تولید کند.

Diffusion می‌تواند نمونه‌های دقیق و با جزئیات بالا تولید کند. چون تعداد زیادی step برای تولید نمونه و همچنین شبکه عصبی قدرتمندی برای فرآیند reverse دارد. عکس‌های تولید شده توسط diffusion از تنوع و کیفیت بیشتری نسبت به VAE و GAN برخوردار است.

کاربرد :

از کاربردهای مدل‌های generative می‌توان ساخت عکس ، افزودن data (augmentation) تشخیص anomaly را نام برد.

تولید و ساخت عکس توسط diffusion و VAE تنوع بیشتری دارد. امّا GAN دارای تنوع کمتر امّا نمونه‌های باکیفیت می‌باشد.

از GAN و diffusion در تسک image-image translation استفاده می‌شود که برای تبدیل یک فرمت عکس به فرمت دیگر است مثل domain transform یا style transfer

همچنین چون هر سه مدل قابلیت بازتولید عکس‌های جدید دارند می‌توانند به عنوان data aug استفاده شوند.

برای تشخیص outlier یا anomaly data هم می‌توان از VAE استفاده کرد که این کار را با محاسبه reconstruction loss انجام می‌دهد یا از GAN و Diffusion استفاده کرد که این کار را به ترتیب با اندازه‌گیری امتیاز discriminator و امتیاز reverse diffusion انجام می‌دهند.

ادامه سوال ①

2. مدل های Diffusion در تسک های مختلفی مثل denoising، Segmentation، image-image translation در تصاویر پزشکی استفاده می شوند.

دلایل استفاده از آن ها عبارتند از: (مزیت ها)

- توانایی هندل کردن دیتا های پیچیده با ابعاد بالا مثل تصاویر پزشکی 3D
- با داده‌های برچسب دار و همچنین غیر برچسب دار آموزش داده می شوند و چون در حوزه پزشکی داده‌های با برچسب محدود است مناسب است.
- توانایی تولید عکس هایی با کیفیت و جزئیات بالا و همچنین حفظ distribution داده ها را دارند.
- می توانند از دانش های قبل مثل segmentation mask و اطلاعات کلینیکی نیز استفاده کنند و با آن فرایند تولید را انجام دهد.

سوال ②

1. چون داده های دیتاست ما یک mani fold ای از فضای IR را اشغال می‌کنند و اگر به داده‌ها نویز اضافه نکنیم تا پخش کند و بیشتر از mani fold را اشغال کنند، گرادیان هایمان در نقطه initial، بی‌معنی (non sense) خواهد بود و به سمت توزیع درست داده ها، هدایت نمی‌شویم. پس نیاز به فرایند دیفیوژن داریم تا در هر مرحله یک نویز کم به دیتاست اضافه کنیم تا $\nabla_x P(x)$ مقدار معنی داری پیدا کرده و ما را به توزیع اصلی نزدیک کند.

$$P(x,t) = \sum_{j=1}^{M} w_j \frac{1}{\sqrt{2\pi\,(s_j^2+\sigma^2 t)}} \exp\left\{-\frac{[x-\mu_j]^2}{2(s_j^2+\sigma^2 t)}\right\}^{\cdot 2}$$

$$\nabla_x P(x,t) = \sum_{j=1}^{M} w_j \frac{1}{\sqrt{2\pi(s_j^2+\sigma^2 t)}} \cdot e^{\left[-\frac{(x-\mu_j)^2}{2(s_j^2+\sigma^2 t)}\right]} \cdot \frac{-(x-\mu_j)}{s_j^2+\sigma^2 t}$$

$$\nabla_x \log P(x,t) = \frac{\nabla_x P(x,t)}{P(x,t)}$$

$$\nabla_x \log P(x,t) = \frac{\sum_{j=1}^{M} w_j \frac{1}{\sqrt{2\pi(s_j^2+\sigma^2 t)}} e^{\left[-\frac{(x-\mu_j)^2}{2(s_j^2+\sigma^2 t)}\right]} \cdot \frac{-(x-\mu_j)}{s_j^2+\sigma^2 t}}{\sum_{j=1}^{M} w_j \frac{1}{\sqrt{2\pi\,(s_j^2+\sigma^2 t)}} e^{\left[-\frac{(x-\mu_j)^2}{2\,(s_j^2+\sigma^2 t)}\right]}}$$

$$t=0 \Rightarrow \nabla_x \log P(x,0) = \frac{\sum_{j=1}^{M} \frac{w_j}{s_j\sqrt{2\pi}} e^{\left[-\frac{(x-\mu_j)^2}{2s_j^2}\right]} \cdot \frac{-(x-\mu_j)}{s_j^2}}{\sum_{j=1}^{M} \frac{w_j}{s_j\sqrt{2\pi}} e^{\left[-\frac{(x-\mu_j)^2}{2s_j^2}\right]}}$$

ادامه سوال ②

بخش 2. به ازای $t=0$ می‌بینیم که ترم $\sigma^2$ در تابع score از بین می‌رود و تاثیرگذار نیست. امّا به ازای $t$ دلخواه غیر صفر با scale $t$ ترم $\sigma^2$ را در تابع score داریم و تاثیرگذار است.

سوال ③

۱. بطور مثال اگر روی تسک تشخیص وضعیت سرطان کلیه کار می‌کنیم، می‌توانیم 100 نمونه اشتباه کلاس‌بندی شده را تحلیل کنیم. مثلا اگر تحلیل بصورت زیر باشد:

| | benign | malignant | noisy |
|---|---|---|---|
| 1 | ✓ | ✓ | |
| 2 | | | |
| 3 | ✓ | ✓ | ✓ |
| ⋮ | | | |
| 100 | | ✓ | |
| | 7% | 63% | 6% |

total dev error 15%
benign ~ 1%
malignant ~ 9.5%
noisy ~ 0.9%

خب در اینصورت وقت گذاشتن برای تهیه داده‌های غیرنویزی یا denoise کردن آن‌ها با یک شبکه عصبی دیگر با توجه به اینکه 0.9% بهبود در error دیتاست dev خواهد داشت، بیهوده است.

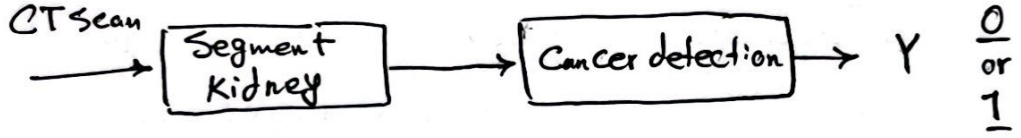
بهتر است علت error زیاد روی کلاس malignant را بررسی کنیم و با افزایش داده‌های مربوط به این کلاس یا تنظیم در loss، بهبود در خطای این کلاس و بهبود قابل توجه (حدود 9.5%) در خطای کل دیتاست dev بوجود بیاوریم.

2. اگر دیتاست dev ما به اندازه کافی بزرگ باشد می‌توانیم devset را به دو مجموعه eyeball devset و blackbox devset تقسیم کنیم. eyeball مجموعه کوچکی از devset است که اجازه داریم به آن با چشم نگاه کنیم و error analysis را روی آن انجام دهیم.

تقسیم devset به این دو مجموعه به این دلیل است که ما سعی می‌کنیم که مدل خودمان را روی eyeball، overfit کنیم و بعد از انجام error analysis و بهبود مدل اگر بخواهیم مدل را دوباره ارزیابی کنیم احتیاج به داده‌های labeled جدید یا مجموعه blackbox داریم. روی مجموعه blackbox در این زمان کار بردی است امّا فقط اجازه داریم روی این مجموعه ارزیابی کرده، میزان خطا را محاسبه کنیم و با توجه به آن هایپرپارامترها را set کنیم ولی اجازه دیدن داده‌های این مجموعه را نداریم.

## سوال ③

3. بطور مثال مانند بخش اول می‌خواهیم داده‌های CT scan را به دو کلاس دارای توده کلیه و فاقد توده کلیوی تقسیم بندی کنیم. برای این کار یک سیستم طراحی می‌کنیم که دارای دو بخش است. بخش اول محدوده کلیه را در تصاویر CT سگمنت می‌کند. بخش دوم محدوده کلیه را گرفته و وجود توده را classify می‌کند.




بوسیله روش error analysis by part می‌توان هر خطا در خروجی (Y) را به یک بخش از سیستم که مقصر بوده، نسبت بدهیم.

بطور مثال برای یک داده خروجی 0 (فاقد سرطان) تشخیص داده شده در حالی که سرطان داشته است. باید برای این مورد بررسی کنیم خروجی شبکه Segmentation چه بوده است. اگر شبکه Segmentation، محدوده کلیه را درست شناسایی نکرده باشد و عضو دیگری را Segment کرده باشد پس شبکه Cancer detection درست کار کرده و خطا مربوط به بخش Segmentation است.

همینطور برای خطاهای مختلف سیستم این تحلیل را انجام می‌دهیم تا ببینیم کدام بخش از سیستم تاثیر بیشتری در خروجی اشتباه داشته است و سپس در اولویت بالاتر باید به تصحیح آن بخش از سیستم بپردازیم.

سوال ④

$$J_{naive}(\theta) = \frac{1}{2} \int dx\, dt\; p(x,t) \left[ S_\theta(x,t) - \nabla_x \log p(x,t) \right]^2$$

base on bayes theorem:

$$p(x,t) = p\left(x,t \middle| x^{(0)}, 0\right) p(x^0, 0)$$

and also:

$$\nabla_x \log p(x,t) = \frac{\nabla_x p(x,t)}{p(x,t)} = \frac{\nabla_x \left[ p\left(x,t \middle| x^{(0)}, 0\right) \cdot p(x^{(0)}, 0) \right]}{p\left(x,t \middle| x^{(0)}, 0\right)\; p(x^{(0)}, 0)}$$

$$\Rightarrow \nabla_x \log p(x,t) = \frac{\left[ \nabla_x p\left(x,t \middle| x^{(0)}, 0\right) \right] \cdot p(x^{(0)}, 0)}{p\left(x,t \middle| x^{(0)}, 0\right)\; p(x^{(0)}, 0)}$$

$$\Rightarrow \nabla_x \log p(x,t) = \frac{\nabla_x p\left(x,t \middle| x^{(0)}, 0\right)}{p\left(x,t \middle| x^{(0)}, 0\right)}$$

$$\Rightarrow \boxed{\nabla_x \log p(x,t) = \nabla_x \log p\left(x,t \middle| x^{(0)}, 0\right)}$$

$$\left[ \frac{\partial}{\partial x} p\left(x,t \middle| x^{(0)}, 0\right) \right] \cdot p(x^{(0)}, 0) + \left[ \frac{\partial}{\partial x} p(x^{(0)}, 0) \right] \cdot p\left(x,t \middle| x^{(0)}, 0\right)$$

صفر

So:

$$J_{naive}(\theta) = \frac{1}{2} \int dx\, dt\; p\left(x,t \middle| x^{(0)}, 0\right) p(x^{(0)}, 0) \left[ S_\theta(x,t) - \nabla_x \log p\left(x,t \middle| x^{(0)}, 0\right) \right]^2$$

$$\Rightarrow \boxed{J_{naive}(\theta) = J_{mod}(\theta)}$$

پس زمانی که دو تابع با هم برابر باشند، global min یکسانی دارند.