



نکات زیر را رعایت کنید:  
 فایل پاسخ را به همراه تمامی کدها در یک فایل فشرده و با عنوان HW2\_STD# در سایت Quera.ir بارگذاری نمایید.  
 بخش‌های پیاده‌سازی مربوط به هر سوال را در فایل مربوطه با شماره‌ی آن سوال قرار دهید.  
 سوالات خود را از طریق Quera مطرح کنید.

## مسئله‌ی ۱. Network Design

(آ) فرض کنید که شبکه‌ای با ساختار زیر داریم، خروجی هر لایه و تعداد پارامترهای آن را محاسبه کنید.

$$\text{Input} = 224 * 224 * 3 \rightarrow 1$$

$$1 \rightarrow \text{Conv}[3 * 3 * 64, \text{stride} = 1, \text{pad} = 1] \rightarrow 2$$

$$2 \rightarrow \text{ReLU} \rightarrow 3$$

$$3 \rightarrow \text{Pooling}[2 * 2 * 32, \text{stride} = 2] \rightarrow 4$$

$$4 \rightarrow \text{FC}[10 - \text{class}] \rightarrow 5$$

\* اعداد ۱ تا ۴ برای مشخص کردن خروجی و ورودی‌های لایه‌ها استفاده شده‌اند.

(ب) در صورتیکه بخواهیم فقط با یک لایه‌ی تمام متصل (FC) تصویری با ابعاد ورودی و خروجی مشابه داشته باشیم، تعداد پارامترهای لایه‌ی تمام متصل را محاسبه کنید

(ج) تفاوت Deformable Convolution با Standard Convolution در چیست و چه مزایایی نسبت به حالت استاندارد دارد؟

## مسئله‌ی ۲. Transposed Convolution

کانولوشن بین ورودی X و فیلتر W به صورت زیر است:

$$X = \begin{bmatrix} x_{(0,0)} & x_{(0,1)} & x_{(0,2)} \\ x_{(1,0)} & x_{(1,1)} & x_{(1,2)} \\ x_{(2,0)} & x_{(2,1)} & x_{(2,2)} \end{bmatrix} \quad W = \begin{bmatrix} w_{(0,0)} & w_{(0,1)} \\ w_{(1,0)} & w_{(1,1)} \end{bmatrix}$$

می‌توان عملیات کانولوشن را به صورت ضرب ماتریسی نوشت که ورودی و خروجی را به صورت یک بردار و فیلتر را به صورت یک ماتریس در نظر گرفت. ورودی X را به صورت بردار زیر نمایش می‌دهیم:

$$X = \begin{bmatrix} x_{(.,0)} & x_{(.,1)} & \dots & x_{(.,r)} \end{bmatrix}^T$$

(آ) عملیات کانولوشن بالا با  $S = 1$  را به صورت ضرب ماتریسی  $Y = AX$  بنویسید.

(ب) با استفاده از نمایش ماتریسی بالا می توان گرادیان پس انتشار نسبت به ورودی  $X$ ، را به صورت  $\frac{\partial L}{\partial X} = A^T \frac{\partial L}{\partial Y}$  نمایش داد.

عملیات transposed convolution را می توان مشابه عملیات گرادیان پس انتشار نسبت به ورودی در نظر گرفت و می توان این عملیات را به صورت کانولوشن مستقیم با ماتریس  $A^T$  به عنوان فیلتر دانست.

فرض کنید  $X$  خروجی یک عملیات کانولوشن با فیلتر  $W$  و  $\text{stride} = 2$  است. خروجی transposed convolution را با ورودی های زیر محاسبه کنید.

$$X = \begin{bmatrix} 2 & 1 \\ 3 & 6 \end{bmatrix} \quad W = \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix}$$

### مسئله ۳. Object Detection

یکی از کاربردهای شبکه های کانولوشن، آشکارسازی اشیا است. مقالات مرتبط را مطالعه کرده و به سوالات زیر پاسخ دهید.

(آ) روش های YOLO و RCNN و RFCN را بررسی کرده و با یکدیگر مقایسه کنید.

(ب) در آشکارسازی های real-time استفاده از کدام یک را پیشنهاد می کنید.

(ج) لایه Spatial Pyramid Pooling را مختصرا توضیح داده و مزایای استفاده از آن را، در آشکارسازی تصاویر بیان کنید.

### مسئله ۴. Classification

در این بخش میخواهیم یک شبکه ی کانولوشنی برای تشخیص تومور مغزی از روی داده های MRI مربوط به مغز طراحی کنیم.

\*\* لطفا تمامی مراحل را به ترتیب انجام دهید و نتایج خواسته شده را گزارش کنید.

\*\* باتوجه به اینکه تعداد ایپاک های آموزشی و پارامترهای شبکه زیاد هستند، پیشنهاد می شود که حتما از Colab در حالت GPU استفاده کنید .

(آ) داده های مربوط به مغز را با استفاده از لینک زیر دانلود کنید . در صورتیکه مایل به استفاده از Colab هستید، میتوانید از راهنمای موجود در لینک قرار داده شده برای لود کردن داده ها استفاده کنید.  
[https://github.com/mralisoltani/CNN\\_Tumor](https://github.com/mralisoltani/CNN_Tumor)

(ب) با استفاده از کد قرار داده شده در فایل Load.py داده ها را در یک DataLoader قرار دهید.

(ج) سائیز و نوع features و label داده های train را چاپ کنید.

(د) یک batch از داده های آموزش را نشان دهید.

(ه) یک شبکه کاملاً متصل به شکل زیر ایجاد کنید:

```
nn.Linear(*, ۵۱۲)
nn.ReLU
nn.Linear(۵۱۲, ۲۴)
nn.ReLU
nn.Linear(۲۴, ۲)
```

- علامت \* در لایه ی اول بدین معنی است که باید سایز ورودی را با توجه به داده ها مشخص کنید و به جای ستاره قرار دهید.

(و) ویژگی های مدل ساخته شده را چاپ کنید

(ز) تابع *train - loop* را برای آموزش مدل، مطابق با شرایط زیر بنویسید:

۱- در این تابع از موارد زیر استفاده کنید:

*model: the neural network*

*dataLoader: a object that fetches training data*

*lossFunction: a function that takes batch-output and batch-labels*

*and computes the loss for the batch*

*optimizer: optimizer for parameters of model*

۲- در هر مرحله از انجام محاسبات *backward*، باید مقدار *loss* و تعداد *batch* های در نظر گرفته شده در آن مرحله نشان داده شوند.

(ح) تابع *test - loop* را برای تست شبکه، مطابق با شرایط زیر بنویسید:

۱- در این تابع از موارد زیر استفاده کنید:

*model: the neural network*

*dataLoader: a object that fetches data*

*lossFunction: a function that takes batch-output and batch-labels*

*and computes the loss for the batch*

۲- در پایان هر مرحله از تست باید مقدار درصد دقت و میانگین *loss* نشان داده شود.

(ط) با در نظر گرفتن مقادیر زیر برای پارامترها، شبکه را برای ۱۰۰ اپاک آموزش دهید.

$LearningRate = 1e - ۳$

$LossFunction = CrossEntropy$

$Optimizer = SGD$

(ی) نمودار درصد دقت و مقدار *loss* شبکه بر روی داده های تست را بر حسب اپاک رسم کنید.

(ک) یک شبکه کانولوشنی به شکل زیر ایجاد کنید:

```
nn.Conv2d(۳, ۳۲, kernel_size = ۳, stride = ۱, padding = ۱)
nn.ReLU()
nn.Conv2d(۳۲, ۶۴, kernel_size = ۳, stride = ۱, padding = ۱)
nn.ReLU()
nn.MaxPool2d(۲, ۲)
nn.Flatten()
nn.Linear(*, ۵۱۲)
nn.ReLU()
nn.Linear(۵۱۲, ۲۴)
nn.ReLU()
nn.Linear(۲۴, ۲)
```

- علامت \* در لایه ی اول بدین معنی است که باید سائز ورودی را با توجه به داده ها مشخص کنید و به جای ستاره قرار دهید.

(ل) ویژگی های مدل ساخته شده را چاپ کنید

(م) این شبکه را نیز با همان پارامترهای قبلی برای ۱۰۰ اپیاک آموزش دهید.

(ن) نمودار درصد دقت و مقدار loss شبکه را بر روی داده های تست برحسب اپیاک رسم کنید.

(س) چه تفاوتی بین نمودار رسم شده برای مدل ۱ و ۲ وجود دارد؟ چرا؟

## مسئله ۵. Semantic Segmentation

یکی از کاربرد های شبکه UNet در بخش بندی معنایی است. با کامل کردن نوت بوک داده شده، یک شبکه UNet پیاده سازی کرده که با دریافت تصویر سیتی اسکن ریه ی بیمار کرونایی، لکه های درگیری ریه (ground glass opacity) را نشان دهد.

(آ) از معیار IoU (Intersection over Union) برای سنجش صحت مدل استفاده کنید. این معیار برای سنجش میزان همپوشانی دو ناحیه استفاده می شود. هر چه میزان همپوشانی دو ناحیه بیشتر باشد، IoU نیز بزرگتر است.

(ب) روند خطا و دقت شبکه را کامل گزارش کنید.

(ج) خروجی شبکه را برای چندین تصویر از داده های تست، پس از اعمال  $threshold = 0.5$  به همراه ماسک آن نشان دهید.

(د) چه ایده ای برای بهبود مدل دارید در صورتی که با استدلال درست همراه باشد، متناسب با میزان بهبود مدل نمره امتیازی دریافت خواهید کرد. (نمره امتیازی)

\* توجه داشته باشید که برای پیاده سازی مدل تنها مجاز به استفاده از کتابخانه pyTorch هستید.