

به نام خدا



دانشگاه تهران  
دانشکده فنی  
دانشکده برق و کامپیوتر



درس تحلیل و طراحی شبکه های عصبی عمیق

تمرین سوم

آذر ۱۴۰۲

هدف از انجام این تمرین آشنایی با یادگیری لایه به لایه و بخش بندی تصاویر است. در سوال یک با یادگیری لایه به لایه آشنا می شوید و پارامترهای موثر در یادگیری لایه ای را بررسی می کنید. در سوال دوم با بخش بندی تصاویر و معماری های مختلف موجود در این حوزه آشنا می شوید.

**نکته مهم :** در این تمرین باید تمام پیاده سازی ها از جمله معماری ها باید توسط خود شما انجام شود.

**نکته مهم :** به تمام نکات گفته شده در انتهای هر تمرین توجه داشته باشید.

## سوال یک) یادگیری لایه به لایه

هدف این سوال طراحی معماری VGG11 و آموزش آن به دو روش یادگیری انتها به انتها و یادگیری لایه ای است. در روش یادگیری لایه ای شما باید تاثیر پارامترهای مختلف را بررسی کنید.

جدول ۱ - جزئیات مربوط به تعداد فیلترها و سایز کرنلها در ساختار شبکه VGG11

ID	Layer	Kernel Size	Num. Filters	Output Shape
0	Conv2D	3×3	64	(32×32×64)
1	BatchNorm2d	--	--	(32×32×64)
2	ReLU	--	--	(32×32×64)
3	MaxPool2d	2×2	--	(16×16×64)
4	Conv2d	3×3	128	(16×16×128)
5	BatchNorm2d	--	--	(16×16×128)
6	ReLU	--	--	(16×16×128)
7	MaxPool2d	2×2	--	(8×8×128)
8	Conv2d	3×3	256	(8×8×256)
9	BatchNorm2d	--	--	(8×8×256)
10	ReLU	--	--	(8×8×256)
11	Conv2d	3×3	256	(8×8×256)
12	BatchNorm2d	--	--	(8×8×256)
13	ReLU	--	--	(8×8×256)
14	MaxPool2d	2×2	--	(4×4×256)
15	Conv2d	3×3	512	(4×4×512)
16	BatchNorm2d	--	--	(4×4×512)
17	ReLU	--	--	(4×4×512)
18	Conv2d	3×3	512	(4×4×512)
19	BatchNorm2d	--	--	(4×4×512)
20	ReLU	--	--	(4×4×512)
21	MaxPool2d	2×2	--	(2×2×512)
22	Conv2d	3×3	512	(2×2×512)
23	BatchNorm2d	--	--	(2×2×512)
24	ReLU	--	--	(2×2×512)

25	Conv2d	3×3	512	(2×2×512)
26	BatchNorm2d	--	--	(2×2×512)
27	ReLU	--	--	(2×2×512)
28	MaxPool2d	2×2	--	(1×1×512)
29	Dense	--	--	512
30	Dense	--	--	10

### نکات مربوط به گزارش نتایج در کل سوال (۱) :

انتظار می رود در این سوال به ازای هر آموزش انجام شده (در تمام بخش های سوال) بر روی شبکه VGG11 موارد خواسته شده زیر را گزارش نمائید.

- نمودار دقت<sup>۱</sup> و تابع هزینه<sup>۲</sup> داده های آموزشی<sup>۳</sup> و ارزیابی<sup>۴</sup> در حین آموزش شبکه در یادگیری انتهای **به انتها**
- نمودار دقت<sup>۵</sup> داده های آموزشی<sup>۶</sup> و ارزیابی<sup>۷</sup> در حین آموزش شبکه برای هر لایه (بعد از اتمام آموزش هر لایه در یادگیری لایه به لایه)
- دقت شبکه بر روی داده های تست بهترین نسخه از شبکه آموزش داده شده
- نمودار مقدار Center SI داده های آموزشی و تست در تمام لایه ها بعد از اتمام آموزش شبکه بر روی بهترین نسخه از شبکه آموزش داده شده
- تحلیل نتایج هر قسمت سوال

### نکاتی در ارتباط با مجموعه داده در کل سوال (۱) :

- از داده های تست اختصاص داده شده صرفاً برای تست نهایی مدل در انتهای آموزش شبکه استفاده کنید و داده های ارزیابی را از داده های آموزشی جدا نمائید. برای این تمرین دقیقاً ده درصد داده ها را جدا نمائید.

---

<sup>1</sup> Accuracy  
<sup>2</sup> Loss Function  
<sup>3</sup> Train  
<sup>4</sup> Validation  
<sup>5</sup> Accuracy  
<sup>6</sup> Train  
<sup>7</sup> Validation

- با توجه به متوازن<sup>۱</sup> بودن مجموعه داده‌ها، لازم است داده‌های آموزشی و ارزیابی نیز به صورت متوازن انتخاب بشوند.
- پیش از استفاده از مجموعه داده در آموزش و ارزیابی شبکه، حتما داده‌ها را نرمال کنید.
- از روش‌های تقویت داده مناسب جهت افزایش دقت شبکه و جهت جلوگیری از Overfitting استفاده نمائید.
- از تمام داده‌ها برای آموزش، ارزیابی و تست شبکه استفاده کنید و داده‌ای را کنار نگذارید.
- از ده درصد داده‌ها برای محاسبه شاخص SI استفاده کنید.

### نکاتی در ارتباط با شبکه عصبی در کل سوال (۱) :

- معماری پیاده‌سازی شده عینا معادل جدول (۱) باشد.
- برای ابر پارامترهای مورد استفاده در آموزش شبکه از پارامترهای گفته شده در جدول (۲) استفاده نمائید.
- وزن‌های اولیه مدل باید به صورت رندوم باشد.
- انتظار می‌رود دقت مدل بر روی مجموعه داده CIFAR10 در روش انتها به انتها بالای ۸۵ درصد باشد.

Parameter	Value
<b>Optimizer</b>	SGD with Momentum
<b>Momentum</b>	0.9
<b>Weight Decay</b>	5e-4
<b>Loss Function</b>	Categorical Cross Entropy
<b>Batch Size</b>	128
<b>Epochs (End-to-End Learning)</b>	200
<b>Epochs Per Stage (Layer-wise Learning)</b>	15
<b>Learning Rate Scheduler</b>	StepLR
<b>Initial Learning Rate</b>	با استفاده از جستجو مقدار اولیه نرخ یادگیری و مقدار Step را پیدا کنید

جدول ۲ - ابرپارامترهای مورد استفاده در سوال (۱)

<sup>۱</sup> Balanced

Number	Subject
1	RandomHorizontalFlip()
2	RandomRotation(15)
3	ColorJitter(brightness=0.2, contrast=0.2, saturation=0.2),

جدول ۳ - روش های تقویت داده مورد استفاده در سوال (۱)

**الف)** یک شبکه VGG11 مطابق جدول (۱) طراحی کنید و با استفاده از پارامترهای جدول (۲) بر روی مجموعه داده CIFAR10 آموزش بدهید. دقت داشته باشید موارد گفته شده در ابتدای سوال را برای آن گزارش نمایید. همچنین روش های تقویت داده گفته شده در جدول (۳) را استفاده نمایید.

**ب)** شبکه VGG11 را به صورت لایه ای آموزش دهید. در آموزش لایه ای، لایه های کانولوشنی را یک به یک اضافه می کنید و پس از منجمد کردن لایه های پیشین، دو لایه Fully connected به شبکه افزوده و با تابع هزینه گفته شده در جدول (۲) شبکه را آموزش دهید. (راهنمایی: سایر لایه ها یعنی maxpool و batch normalization را همراه با لایه های کانولوشنی اضافه کنید و نیاز به در نظر گرفتن یک Stage جداگانه برای آنها نیست).

**مثال:** فرض کنید ساختار شبکه به صورت زیر است:

[conv, relu, conv, relu, conv, relu, maxpool]

برای آموزش لایه ای شبکه، باید طبق ترتیب زیر لایه ها را اضافه کنید:

L1 = [conv, relu]

L2 = [conv, relu]

L3 = [conv, relu, maxpool]

Stage 1:

Net = [L1, FC]

Train Net (L2, FC=Train)

Stage 2:

Freeze L1

Net = [L1, L2, FC]

Train Net (L1=Freeze, L2, FC=Train)

Stage 3:

Freeze L1, L2

Net = [L1, L2, L3, FC]

Train Net (L1, L2=Freeze, L3, FC=Train)

**نکته مهم :** در مجموع نه Stage برای یادگیری لایه به لایه نیاز دارید. با توجه به جدول (۲) برای هر Stage نیز به ۱۵ اپاک برای یادگیری نیاز دارید که مجموعاً برابر با ۱۳۵ دوره آموزشی خواهد بود.

(ج) دقیقاً مانند قسمت (ب) عمل کنید؛ ولی لایه‌های قبلی را منجمد نکنید. نتایج را با قسمت‌های قبل مقایسه کنید.

**مثال:** فرض کنید ساختار شبکه به صورت زیر است:

[conv, relu, conv, relu, conv, relu, maxpool]

برای آموزش لایه ای شبکه ، باید طبق ترتیب زیر لایه‌ها را اضافه کنید:

```
L1 = [conv, relu]
L2 = [conv, relu]
L3 = [conv, relu, maxpool]
Stage 1:
    Net = [L1, FC]
    Train Net (L1,FC=Train)

Stage 2:
    Freeze L1
    Net = [L1, L2, FC]
    Train Net (L1, L2,FC=Train)

Stage 3:
    Freeze L1, L2
    Net = [L1, L2, L3, FC]
    Train Net (L1,L2, L3,FC=Train)
```

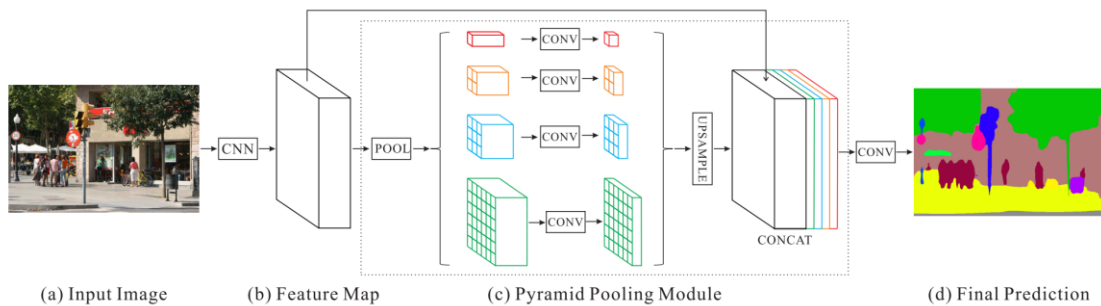
**نکته مهم :** در مجموع نه مرحله برای یادگیری لایه به لایه نیاز دارید. با توجه به جدول (۱) برای هر Stage نیز به ۱۵ اپاک برای یادگیری نیاز دارید که مجموعاً برابر با ۱۳۵ دوره آموزشی خواهد بود.

(د) دقیقاً مانند قسمت (ج) عمل کنید؛ با این تفاوت که در چهار مرحله ابتدایی از تمام روش های تقویت داده گفته شده در جدول (۳) برای آموزش استفاده کنید و در باقی مراحل تنها از یک روش به انتخاب خودتان استفاده کنید.

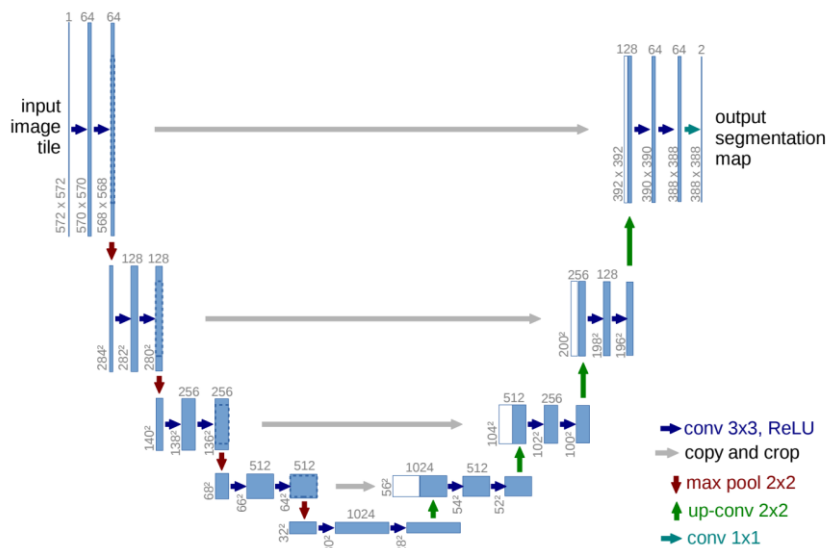
(و) دقیقاً مانند قسمت (ج) عمل کنید؛ با این تفاوت که در چهار مرحله ابتدایی تنها از یک روش تقویت داده از روش های گفته شده در جدول (۳) برای آموزش استفاده کنید و در باقی مراحل از همه روش های تقویت داده استفاده کنید.

## سوال دو) شبکه بخش بندی تصاویر دو بعدی

هدف از این سوال، پیاده‌سازی یک شبکه برای بخش بندی تصاویر است. در مباحث درس با شبکه های بخش بندی تصویر آشنا شده‌اید. هدف این تمرین پیاده سازی دو شبکه UNet<sup>1</sup> و PSPNet<sup>2</sup> است.



شکل (۱) - معماری PSPNet



شکل (۲) - معماری UNet

<sup>1</sup> Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015:

<sup>2</sup> Zhao, Hengshuang, et al. "Pyramid scene parsing network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.



مجموعه داده SUIM<sup>۱</sup> مجموعه داده ای شامل تصاویر زیر آب و بخش بندی اجسام و موجودات زیر آب است که نمونه ای از تصاویر آن را در شکل (۳) مشاهده می نمائید. مشخصات این مجموعه داده را از اینجا<sup>۲</sup> می توانید مشاهده نمائید و موارد مورد نیاز را دانلود نمائید.



شکل (۳) - مجموعه داده SUIM

### نکات مربوط به گزارش نتایج در کل سوال (۱) :

انتظار می رود در این سوال به ازای هر آموزش انجام شده (در تمام بخش های سوال) بر روی شبکه VGG11 موارد خواسته شده زیر را گزارش نمائید.

- نمودار mIOU و تابع هزینه<sup>۳</sup> داده های آموزشی<sup>۴</sup> و ارزیابی<sup>۵</sup> در حین آموزش شبکه
- مقدار mIOU شبکه بر روی داده های تست بهترین نسخه از شبکه آموزش داده شده
- نمونه تصاویر (۱۰ تصویر) استخراج شده از شبکه مانند تصاویر بخش بندی شده شکل (۳)

### نکاتی در ارتباط با مجموعه داده در کل سوال (۱) :

- از داده های تست اختصاص داده شده صرفاً برای تست نهایی مدل در انتهای آموزش شبکه استفاده کنید و داده های ارزیابی را از داده های آموزشی جدا نمائید. برای این تمرین دقیقاً ده درصد داده ها را جدا نمائید.
- لازم است داده های آموزشی و ارزیابی تا حد ممکن به صورت متوازن انتخاب بشوند و از هر دسته به میزان درصد برابر از هر برچسب در داده های آموزشی و ارزیابی وجود داشته باشد.
- پیش از استفاده از مجموعه داده در آموزش و ارزیابی شبکه، حتماً داده ها را نرمال کنید.

<sup>۱</sup> Islam, Md Jahidul, et al. "Semantic segmentation of underwater imagery: Dataset and benchmark." 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2020.

<sup>۲</sup> <https://irvlab.cs.umn.edu/resources/suim-dataset>

<sup>۳</sup> Loss Function

<sup>۴</sup> Train

<sup>۵</sup> Validation

- از روش های تقویت داده مناسب جهت افزایش دقت شبکه و جهت جلوگیری از Overfitting استفاده نمائید.
- از تمام داده ها برای آموزش، ارزیابی و تست شبکه استفاده کنید و داده ای را کنار نگذارید.

### نکاتی در ارتباط با شبکه عصبی در کل سوال (۱) :

- از ابعاد ورودی  $۱۶۰ \times ۱۶۰$  یا  $۱۶۰ \times ۱۲۰$  برای آموزش مدل استفاده کنید.
- از ابرپارامترهای مناسب هر مدل بر اساس مقاله استفاده نمائید.

**الف)** سازوکار و معماری دو شبکه UNet و PSNet را از روی مقالات آنها مطالعه کنید و هر کدام را مختصراً توضیح دهید.

**ب)** شبکه UNet را پیاده سازی نمائید و مجموعه داده SUIM را بر روی آن آموزش بدهید و موارد خواسته شده در ابتدای سوال را برای آن گزارش نمائید.

**ج)** شبکه PSPnet را با Backbone شبکه MobileNetV2 پیاده سازی نمائید و مجموعه داده SUIM را بر روی آن آموزش بدهید و موارد خواسته شده در ابتدای سوال را برای آن گزارش نمائید.

## نکات:

لطفاً نکات گفته شده را به دقت مطالعه نمائید، در صورت عدم رعایت هر کدام از موارد گفته شده نمره کسر خواهد شد.

- مهلت تحویل این تمرین، **پنج شنبه ۷ دی** است.
- مهلت تحویل تمرین **قابل تمدید نیست**.
- انجام این تمرین به صورت **یک یا دو نفره** می باشد.
- گزارش شما در فرآیند تصحیح از اهمیت ویژه ای برخوردار است. لطفاً تمامی نکات و فرض هایی که برای پیاده سازی ها و محاسبات خود در نظر می گیرید را در گزارش ذکر کنید.
- **کدهای** خود را به صورت **عکس** در داخل گزارش **کپی نکنید** و با فرمتی مناسب آن را در گزارش قرار دهید.
- داخل کدها **کامنت** های لازم را قرار دهید و تمامی موارد مورد نیاز برای اجرای صحیح کد را ارسال کنید.
- الزامی به ارائه توضیح جزئیات کد در گزارش نیست. اما باید نتایج بدست آمده را گزارش و تحلیل کنید.
- گزارش را در قالب تهیه شده که روی صفحه درس در سامانه eLearn بارگذاری شده، بنویسید. در صورت تمایل می توانید از Latex نیز برای نوشتن گزارش استفاده نمائید اما باید ساختار، زبان نوشتار و سایر موارد قالب اصلی را نیز رعایت کنید.
- در گزارش خود برای **تصاویر زیرنویس** و برای **جداول هم بالانویس** اضافه کنید.
- اگر بخشی از کد را از کدهای آماده اینترنتی استفاده می کنید که جزء قسمتهای اصلی تمرین نمی باشد، حتماً باید لینک آن در گزارش و کد ارجاع داده شود، در غیر اینصورت تقلب محسوب شده و کل نمره تمرین را از دست می دهید. ولی محدودیتی در استفاده از منابع اینترنتی ندارید و در مواردی که در تمرین اشاره نشده می توانید از کدهای موجود استفاده کنید.
- تنها مجاز به استفاده از زبان برنامه نویسی Python و یکی از دو کتابخانه Tensorflow یا Pytorch برای پیاده سازی شبکه های عصبی هستید.
- هر گونه **شباهت** در گزارش و کدها، به منزله **تقلب** می باشد و نمره تمامی افراد شرکت کننده در آن تقلب برابر با صفر خواهد بود.
- لطفاً فایل گزارش، کدها و سایر ضامین مورد نیاز را با فرمت زیر در صفحه درس در سامانه eLearn بارگذاری نمائید.

- HW3\_[Lastname1]\_[StudentNumber1]\_[Lastname2]\_[StudentNumber2]..zip

Or

- HW3\_[Lastname]\_[StudentNumber].zip

به طور مثال:

HW3\_Karimi\_8101234567\_Ahmadi\_810135345.zip

- در صورت وجود هرگونه ابهام یا مشکل می‌توانید از طریق رایانامه زیر با دستیار آموزشی طراح تمرین در تماس باشید:

سوال اول (علی رشیدی) :

[Alirashidimoghadam@gmail.com](mailto:Alirashidimoghadam@gmail.com)

سوال دوم (علی کریمی) :

[alikarimi120@gmail.com](mailto:alikarimi120@gmail.com)