

به نام خدا



دانشگاه تهران  
پردیس دانشکده‌های فنی  
دانشکده برق و کامپیوتر



## شبکه‌های عصبی مصنوعی و یادگیری عمیق

مینی پروژه شماره ۱

آبان ۹۹

## فهرست سوالات

سوال ۱ - CNN	۳
سوال ۲ - Deep Learning	۴
سوال ۳ - Object Detection	۷
سوال ۴ - Semantic Segmentation	۸

این سوال در ادامه سوال دوم تمرین دوم طرح شده است ، هدف در این تمرین افزایش دقت شبکه طبقه بند تصویر با اضافه کردن لایه های کانولوشنی و .. به شبکه است . در این سوال از کد سوال دوم تمرین دوم استفاده کنید و آن را توسعه دهید ( اگر مینی پروژه را به صورت گروهی انجام می دهید می توانید کد یکی از اعضای گروه در تمرین دوم را به صورت مشترک استفاده نمایید ) .

جهت یادآوری ، حتما موارد زیر را در گزارش برای قسمت های (الف) تا (د) بیاورید :

- در دو نمودار جداگانه تغییرات دقت<sup>۱</sup> و خطای آمدل در هر دور<sup>۲</sup> را برای داده ی ارزیابی و آموزش حالت های خواسته شده نشان دهید.
- دقت و خطای نهایی برای داده های تست را نیز گزارش نمایید.

الف ) به بهترین شبکه بدست آمده در سوال دوم تمرین دوم ، لایه های کانولوشنی را اضافه نمایید و شبکه را پیاده سازی نمایید. نتیجه بدست آمده را از نظر دقت و خطا با معماری MLP مقایسه نمایید.

ب ) لایه های Batch normalization و Pooling را توضیح دهید و سپس این لایه ها را به توپولوژی شبکه اضافه نمایید و شبکه را پیاده سازی نمایید. نتیجه بدست آمده را از نظر دقت و خطا با معماری قسمت (الف) مقایسه نمایید.

ج ) به معماری شبکه بدست آمده در قسمت (ب) اکنون Dropout را نیز اضافه نمایید و تاثیر آن را بررسی نمایید. چرا از Dropout در معماری شبکه عصبی استفاده می نمائیم؟

د ) توقف زود هنگام در شبکه های عصبی به چه معناست ؟ چه معیارهایی در این توقف زود هنگام استفاده می شوند؟ یک نمونه از آن را پیاده سازی نمایید.

<sup>1</sup> Accuracy

<sup>2</sup> Loss

<sup>3</sup> Epoch

## سوال ۲ – Deep Learning

با پیشرفت های اخیر در استفاده از CNN در حوزه بینایی کامپیوتر ، مدل های معروفی از شبکه های عصبی کانولوشنی بوجود آمدند . در این سوال تصمیم گرفته شده است گروه ها مدل های مختلفی را مطالعه و پیاده سازی نمایند. این تقسیم بندی بر اساس رقم آخر شماره دانشجویی شما انجام شده است.

**نکته :** برای محاسبه رقم آخر اگر مینی پروژه را در قالب گروه ۲ نفره انجام می دهید ، رقم آخر شماره های دانشجویی خود را با هم جمع نموده و باقی مانده آن را بر ۱۰ بگیرید. در صورتی که مینی پروژه را به صورت ۱ نفره انجام می دهید تنها کافی است بر اساس آخرین رقم شماره دانشجویی خود مدل را انتخاب نمایید.

**نکته :** با توجه به محدود بودن نسخه مدل ها در گزینه ۸ و ۹ ، دو مدل در این قسمت ها آمده است. در صورتی که یکی از گزینه های ۸ یا ۹ برای شما انتخاب شد ، لازم است در قسمت الف هر دو مدل را توضیح دهید و در قسمت (ب) تا (ه) فقط یک مدل را انتخاب و خواسته های این قسمت ها را انجام دهید.

مدل	رقم آخر شماره دانشجویی
EfficientNet	۰
MobileNet	۱
Inception	۲
ResNet	۳
DenseNet	۴
ShuffleNet	۵
NASNet	۶
VGG	۷
GoogLeNet, Xception	۸
Squeezenet , AlexNet	۹

الف ( مختصرا موارد زیر را در مورد معماری انتخاب شده برای شبکه توضیح دهید :

- معماری شبکه
- توضیح نسخه های مختلف معماری و تفاوت آنها (در صورت وجود نسخه های مختلف)
- مزایا و معایب نسبت به سایر مدل ها
- محدودیت های سائز تصویر ورودی
- پیش پردازش های اولیه برای تصویر ورودی
- سائز خروجی مدل و معنای آن

ب ( شبکه انتخابی را با کمک Transfer learning پیاده سازی نمائید و سپس قسمت (ج) را اجرا نمائید.

ج ( ابتدا یک عکس رنگی با کیفیت مناسب از محیط اطراف خود بگیرید. ( این عکس بایستی از دسته های قابل تشخیص توسط مدل باشد! این عکس را در صورت نیاز پیش پردازش کنید، سپس به شبکه ای که پیاده سازی کرده اید داده و ۳ دسته به ترتیب با بیشترین احتمال پیش بینی شده را در خروجی نشان دهید.

د ( اگر تصویر ورودی به شبکه جز اشیا قابل تشخیص توسط مدل نباشد ، خروجی مدل قابل استناد نیست ، راه حل شما برای حل این مشکل چیست؟ آن را پیاده سازی نمائید و نتیجه آن را گزارش نمائید.

ه ( یک لیست از اشیا قابل تشخیص توسط مدل بدست بیاورید ، یک مجموعه داده جمع آوری کنید که شامل دو کلاس باشد ( حداقل هر کلاس شامل ۲۰۰ تصویر باشد، قاعدتا می توانید مجموعه داده بیشتری جمع آوری کنید و محدودیتی وجود ندارد) و مدل را با استفاده از این مجموعه داده مجددا آموزش دهید.

**نکته :** لازم است بنا به مدل انتخاب شده پیش پردازش های لازم روی مجموعه داده را انجام دهید و سپس با استفاده از وزن های مدل که قبلا آموزش داده است ، مجددا شبکه را آموزش دهید.

**نکته :** اینکه چه دسته هایی جمع آوری می کنید و موضوع این دو دسته چیست کاملا بر عهده شماست اما تصاویر انتخابی باید مناسب طبقه بندی باشد. همچنین می توانید از دو دسته از دیتاست های معروف این حوزه همچون ImageNet نیز استفاده کنید.

**نکته:** دیتاست را در سامانه درس آپلود نکنید (فقط برای استفاده خودتان است). ولی توضیحاتی در ارتباط با دو دسته جمع آوری شده در گزارش بیاورید.

موارد زیر را در گزارش برای قسمت (ه) بیاورید:

- در دو نمودار جداگانه تغییرات دقت<sup>۱</sup> و خطای<sup>۲</sup> مدل در هر دور<sup>۳</sup> را برای داده‌ی ارزیابی و آموزش نشان دهید.

---

<sup>۱</sup> Accuracy

<sup>۲</sup> Loss

<sup>۳</sup> Epoch

## سوال ۳ – Object Detection

در بین مدل‌های تشخیص شیء، معماری‌های یک‌پارچه YOLO از لحاظ سرعت و دقت به نتایج چشم‌گیری رسیدند. این الگوریتم‌ها می‌توانند در زمان کمتری در مرحله آموزش به همگرایی برسند و در مرحله استنتاج، اشیاء را با سرعت بالایی تشخیص داده و دسته‌بندی کنند. با استفاده از مطالبی که در ویدئوهای ضبط‌شده YOLO در اختیار شما قرار گرفته است و نیز منابع موجود در اینترنت به سوالات زیر پاسخ دهید.

الف) فرض کنید می‌خواهیم شبکه‌های YOLO-v1 و YOLO-v3 را روی داده‌هایی که دارای ۳۰۰ کلاس هستند آموزش دهیم. لایه آخر هر کدام از این شبکه‌ها دارای چه عمقی<sup>۱</sup> باید باشند؟ (تعداد پهنه‌های پیش‌بینی<sup>۲</sup> را برابر با مقادیر ذکر شده در مقاله‌های اصلی شبکه در نظر بگیرید).

ب) عملکرد YOLO-v1 روی داده‌هایی که دارای اشیائی با تداخل<sup>۳</sup> بالا هستند، چه نقص عمده‌ای دارد؟ توضیح دهید. برای حل این مشکل چه روشی در YOLO-v2 اتخاذ شده است.

ج) در الگوریتم YOLO-v3 برای تشخیص صحیح‌تر داده‌هایی که دارای اشیائی با پراکندگی اندازه‌های بالا<sup>۴</sup> هستند، چه روش‌هایی پیشنهاد شده است؟ توضیح دهید.

د) داده‌های chess\_pieces.zip در اختیار شما قرار گرفته است. با استفاده از این داده‌ها شبکه YOLO-v5 را آموزش دهید<sup>۵</sup>. پس از اتمام فرآیند آموزش، عملکرد شبکه را روی داده‌های تست موجود در دیتاست را با استفاده از معیار mAP گزارش کنید و در نهایت شبکه را روی تعدادی از داده‌های تست اجرا کنید و تصاویر نتایج حاصل و زمان استنتاج را در گزارش بیاورید.

---

<sup>1</sup> Channel

<sup>2</sup> Bounding box

<sup>3</sup> Overlap

<sup>4</sup> High variation in size

<sup>5</sup> <https://github.com/ultralytics/yolov5>

<sup>6</sup> Fine-tune

## سوال ۴ – Semantic Segmentation

در بین شبکه های معرفی شده در راستای قطعه بندی معنایی تصاویر ، دو مدل DeepLab<sup>۱</sup> و Mask-RCNN<sup>۲</sup> از مشهورترین آنها هستند. نمونه ای از تفاوت این دو مدل را با مدل YOLO معرفی شده در تفاوت خروجی آنها در تصویر زیر مشاهده کنید :



یکی از این مدل ها را انتخاب کرده و موارد الف تا ج را برای آن مدل انجام دهید.

الف ( مقاله مدل مربوطه را مطالعه کنید و سازوکار مدل را مختصرا توضیح دهید.

ب ( وزن های از پیش آموزش داده شده مدل را از منابع موجود دانلود کنید و آن را بر روی تصویر انتخابی خودتان اجرا و نتیجه را گزارش کنید.

ج ( از مجموعه داده Soccer استفاده نمائید و مدل را در هر دو حالت زیر آموزش دهید و نتایج آن را به ازای هر اپیک گزارش نمائید.

۱- بازآموزش شبکه با استفاده از وزن های مدل از پیش آموزش داده شده و مجموعه داده Soccer

۲- آموزش شبکه بدون استفاده از وزن های قبلی و با استفاده از مجموعه داده Soccer

همچنین در نهایت با استفاده از یکی از تصاویر قسمت تست ، شبکه را مورد ارزیابی قرار دهید.

<sup>1</sup> <https://arxiv.org/abs/1606.00915>

<sup>2</sup> [https://openaccess.thecvf.com/content\\_ICCV\\_2017/papers/He\\_Mask\\_R-CNN\\_ICCV\\_2017\\_paper.pdf](https://openaccess.thecvf.com/content_ICCV_2017/papers/He_Mask_R-CNN_ICCV_2017_paper.pdf)



دیتاست را از این [لینک](#) می توانید دانلود کنید.

**نکته :** این مدل ها معماری های مختلف و ورژن های متفاوتی دارند ، انتخاب معماری یا ورژن آن بر عهده خودتان است و محدودیتی وجود ندارد و صرفا لینک یکی از معماری ها از هر مدل در پاورقی صفحه قبل قرار داده شده است.

## نکات:

- مهلت تحویل این مینی پروژه تا ۳۰ آبان است.
- گزارش را در قالب تهیه شده که روی صفحه درس در Elearn بارگذاری شده، بنویسید.
- گزارش شما در فرآیند تصحیح از اهمیت ویژه‌ای برخوردار است. لطفاً تمامی نکات و فرض‌هایی که برای پیاده‌سازی‌ها و محاسبات خود در نظر می‌گیرید را در گزارش ذکر کنید.
- در گزارش خود برای تصاویر زیرنویس و برای جداول هم بالانویس اضافه کنید.
- الزامی به ارائه توضیح جزئیات کد در گزارش نیست. اما باید نتایج بدست آمده را گزارش و تحلیل کنید.
- دستیاران آموزشی ملزم به اجرا کردن کدهای شما نیستند. بنابراین هرگونه نتیجه و یا تحلیلی که در شرح سوال از شما خواسته شده است را به طور واضح و کامل در گزارش بیاورید. در صورت عدم رعایت این مورد، بدیهی است که از نمره تمرین کسر می‌شود.
- در صورت مشاهده تقلب امتیاز تمامی افراد شرکت‌کننده در آن، ۱۰۰- لحاظ می‌شود.
- برای انجام تمرین‌ها و مینی پروژه‌ها، تنها زبان برنامه نویسی مجاز Python است. در این مینی پروژه محدودیتی در استفاده از کتابخانه‌ها ندارید.
- استفاده از کدهای آماده برای تمرین‌ها به هیچ وجه مجاز نیست. اما برای مینی پروژه‌ها فقط برای قسمت‌هایی از کد و به عنوان راهنمایی برای پیاده‌سازی، می‌توانید از کدهای آماده استفاده کنید.
- نحوه محاسبه تاخیر به این شکل است: مهلت ارسال بدون جریمه تا تاریخ اعلام شده و پس از آن به مدت هفت روز تا ۷ آذر بارگذاری ممکن است و در نهایت، پس از بازه تاخیر نمره تکلیف صفر خواهد شد.
- لطفاً گزارش، فایل کدها و سایر ضمایم مورد نیاز را با فرمت زیر در سامانه مدیریت دروس بارگذاری نمایید.

Project1 \_[Lastname]\_[StudentNumber].zip

- در صورت وجود هرگونه ابهام یا مشکل می‌توانید از طریق رایانامه‌های زیر با دستیاران آموزشی مربوطه آقایان علی کریمی (سوال ۱ و ۲ و ۴) ، حسام مجتهدی (سوال ۳) و حسین انتظاری (سوال ۳) در تماس باشید :

[Alikarimi120@gmail.com](mailto:Alikarimi120@gmail.com)

[hsm.moj@gmail.com](mailto:hsm.moj@gmail.com)

[h.entezari78@gmail.com](mailto:h.entezari78@gmail.com)

