

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



شبکه های عصبی مصنوعی و یادگیری عمیق

پروژه شماره ۱

فروردین ۱۴۰۱

فهرست سوالات

- سوال ۱ – CNN (Classification) ۳
- سوال ۲ – Transfer Learning ۴
- سوال ۳ – Segmentation ۶
- سوال ۴ – تشخیص اشیا ۸

سوال ۱ – CNN (Classification)

در این سوال با لایه های کانولوشنی در ساختار شبکه های عصبی برای استخراج ویژگی آشنا می شوید. هدف در این تمرین افزایش دقت شبکه طبقه بند تصویر با اضافه کردن لایه های کانولوشنی و .. به شبکه است. برای بخش طبقه بندی در این سوال می توانید از کد سوال اول 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). دیتاست این تمرین [CIFAR-10](#) است که با استفاده از کتابخانه Keras به صورت زیر قابل دانلود است.

```
from keras.datasets import cifar10
(x_train, y_train), (x_test, y_test) = cifar10.load_data()
```

همچنین میتوان از طریق این [لینک](#) مجموعه داده را دانلود کرد.

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

در دو نمودار جداگانه تغییرات دقت^۱ و خطای^۲ مدل در هر دور^۳ را برای داده های ارزیابی و آموزش حالت های خواسته شده نشان دهید.

دقت و خطای نهایی برای داده های تست را نیز گزارش کنید.

(الف) در این بخش تنها با استفاده از لایه های کانولوشنی در بخش استخراج ویژگی، شبکه را پیاده سازی کنید. دقت و خطای شبکه را پس از آموزش، گزارش کنید.

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

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

(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 کنید.

¹ Accuracy

² Loss

³ Epoch

سوال ۲ – Transfer Learning

هدف از این سوال آشنایی با مدل‌های معروف شبکه‌های عصبی کانولوشنی و مطالعه دقیق و پیاده سازی یکی از این شبکه‌ها به کمک transfer learning است.

ابتدا طبق رقم آخر شماره دانشجویی خود یکی از مدل‌های زیر را انتخاب کنید. (اگر گروهی هستید مدل را براساس یکان جمع رقم‌های آخر شماره دانشجویی خود انتخاب کنید.) و به بخش‌های بعدی سوال پاسخ دهید.

رقم آخر دانشجویی	مدل
0	GoogLeNet
1	ResNet
2	Squeeze Net
3	DenseNet
4	Shuffle Net
5	ENet
6	AlexNet
7	VGG- 16
8	VGG- 19
9	Inception

الف) موارد زیر را در مورد مدل انتخابی خود توضیح دهید:

معماری شبکه

مزایا و معایب

در صورتی که تصویر ورودی شبکه به پیش پردازش اولیه ای نیاز دارد آن را توضیح دهید.

ب) تعریفی از transfer learning ارائه دهید و شبکه انتخابی را به کمک transfer learning پیاده سازی کنید و قسمت ج را اجرا کنید.

ج) ابتدا مشخص کنید که چه دسته عکس‌هایی توسط مدل انتخابی شما قابل تشخیص هستند. سپس یک عکس دلخواه از دسته‌های قابل تشخیص را انتخاب کنید. در صورت لزوم پیش پردازش‌های لازم را بر روی عکس انجام دهید و به شبکه‌ای که در بخش قبل طراحی کرده اید دهید و ۳ دسته به ترتیب با بیشترین احتمال پیش بینی شده را در خروجی نشان دهید.

د) یک دسته از عکس های مختلفی که توسط مدل شما قابل تشخیص باشند را به دلخواه انتخاب نمایید، پیش پردازش های لازم را انجام دهید و مدل را دوباره آموزش دهید. (می توانید از دیتاست های معروفی همچون imageNet استفاده نمایید و یا دیتاست های دیگری که به نظر مناسب می آیند)

سوال ۳ - Segmentation

یکی از کاربرد های شبکه عصبی است که طبقه بندی هر پیکسل در یک تصویر از مجموعه ای از کلاس های از پیش تعریف شده را انجام می دهد.



در تصویر نمونه بالا پیکسل های مربوط به تخت خواب با برچسب تخت و پیکسل های مربوط به دیوار یا دیگر اشیاء با برچسب مخصوص خودشان دسته بندی شده اند.

به طور خاص، هدف ما این است که تصویری با اندازه $W \times H \times 3$ بگیریم و یک ماتریس $W \times H$ حاوی شناسه کلاس پیش بینی شده مربوط به تمام پیکسل ها تولید کنیم.



segmented

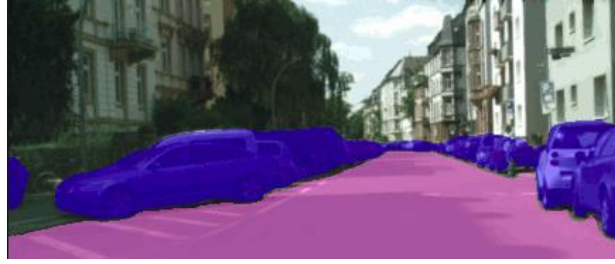
1: Person
2: Purse
3: Plants/Grass
4: Sidewalk
5: Building/Structures

3	3	3	3	3	3	3	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	3	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	1	1	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	1	1	1	3	3	3	5	5	5	5	5	5	5
3	3	3	3	3	1	1	3	3	3	5	5	5	5	5	5	5	5
5	5	3	3	3	3	1	1	3	3	5	5	5	5	5	5	5	5
4	4	3	4	1	1	1	1	1	1	4	4	4	5	5	5	5	5
4	4	3	4	1	1	1	1	1	1	4	4	4	4	4	5	5	5
4	4	4	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	2	2	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	2	2	1	1	1	1	1	4	4	4	4	4	4	4

Input

Semantic Labels

تقسیم بندی معنایی با تشخیص شی متفاوت است، زیرا هیچ جعبه مرزی را در اطراف اشیاء پیش بینی نمی کند و همچنین بین اشیاء در یک دسته تمایز در نظر نمی گیرید بعنوان مثال برای چند خودرو متفاوت در یک تصویر یک برچسب یکسان در نظر می گیرد.



در بین مدل های معروف (FCN-Unet-PSPNet-DEEPLAB,...) ارائه شده دو مدل به دلخواه انتخاب کنید، قبل از پیاده سازی مدل ها معماری هر شبکه را مقایسه کنید و با ذکر دلیل بیان کنید پیش بینی شما از دقت هر شبکه چگونه است سپس با پیاده سازی مدل ها با یک تصویر دلخواه و بررسی خروجی دو مدل را مقایسه کنید.

سوال ۴ – تشخیص اشیا

YOLO یک الگوریتم است که با استفاده از شبکه های عصبی کانولوشن برای تشخیص اشیا مورد استفاده قرار می گیرد. الگوریتم های تشخیص اشیا این امکان را می دهند که علاوه بر تشخیص نوع اشیا مکان آن ها هم در تصویر یا ویدیو را مشخص کنیم. ورژن های مختلفی از این الگوریتم ارائه شده که هر کدام نسبت به ورژن های قبل پیشرفتی ارائه داده اند.

الف) توضیح دهید در YOLOv2 نسبت به YOLOv1 چه ایراداتی برطرف شده و چه پیشرفت هایی در این ورژن جدیدتر نسبت به قبلی صورت گرفت.

ب) در مورد YOLOv5 و YOLOv4 توضیح دهید نسبت به ورژن های قبلی چه تغییراتی داشتند و این تغییرات چگونه در بهبود عملکرد موثر است.

پ) در مبحث تشخیص شیء، شبکه ها به دو دسته ی یک مرحله ای (One-Stage) و دو مرحله ای (Two-Stage) تقسیم میشوند. روند کلی هر کدام از این دو دسته را شرح دهید و برای هر کدام مثالی ذکر کنید.

ت) شبکه YOLOv5 را پیاده سازی کنید. در گزارش بطور خلاصه کد خود را توضیح دهید. و خروجی شبکه را رسم کنید.

راهنمایی: به دلیل محدودیت سخت افزاری برای پیاده سازی شبکه YOLOv5 می توانید از این [دستور العمل](#) و مجموعه داده ای که در اختیار شما قرار گرفته استفاده کنید.

نکات:

- مهلت تحویل این پروژه ۳ اردیبهشت 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.
- 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 ازای هر روز ۵ درصد نمره کسر خواهد شد و حداکثر تا یک هفته امکان ارسال با تاخیر وجود، 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.

MP1_[Lastname]_[StudentNumber].zip

- در صورت وجود هرگونه ابهام یا مشکل می‌توانید از طریق رایانامه‌های زیر با دستیاران آموزشی مربوطه آقایان علی شعاع حقیقی (سوال ۱) و سینا سبزواری (سوالات ۳ و ۴) و خانم شهلا دانشی (سوال ۲) در تماس باشید:

a.shoaahaghi@gmail.com

sinasabzevar@ut.ac.ir

Shahla.daneshi@ut.ac.ir