

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



شبکه های عصبی مصنوعی و یادگیری عمیق

مینی پروژه شماره ۱

پاییز ۱۴۰۰

فهرست سوالات

سوال ۱ - CNN ۳

سوال ۲ - Transfer Learning ۴

سوال ۳ - Semantic Segmentation ۵

سوال ۴ - Object Detection ۶

سوال ۱ – CNN

در این تمرین قصد داریم تا با استفاده از شبکه های CNN مجموعه داده های fashion_mnist را طبقه بندی کنیم.

این مجموعه داده ای از ۶۰۰۰۰ تصویر 28×28 در مقیاس خاکستری از ۱۰ دسته مد به همراه یک مجموعه آزمایشی از ۱۰۰۰۰ تصویر است. با استفاده از کد زیر مجموعه داده را ذخیره کنید.

```
import keras
(X_train, y_train), (X_test, y_test) = keras.datasets.fashion_mnist.load_data()
```

پیش پردازش های لازم را برای آموزش یک شبکه MLP با دو لایه مخفی انجام دهید. در این سوال از روش Stochastic mini batch base استفاده کنید و در هر قسمت نمودار های خطا و دقت و همچنین ماتریس آشفتگی را بدست آورید.

(الف) از سه دسته با اندازه های ۳۲، ۶۴ و ۲۵۶ استفاده نمایید و تاثیرشان را در دقت و خطا بررسی

کنید

(ب) تابع فعالساز و خطا را تغییر داده و تاثیرشان را در دقت و خطا بررسی کنید. (ReLU, TanH و

(MSE, Cross Entropy) (Sig

(ج) به بهترین شبکه بدست آمده لایه های کانولوشنی اضافه کنید نتیجه بدست آمده برای دقت و

خطا را با حالت قبل مقایسه کنید.

(د) عملکرد لایه های Batch normalization و Pooling را توضیح دهید و آنها را به شبکه اضافه

نمایید. سپس نتایج را برای دقت و خطا بررسی کنید.

(ه) لایه Dropout چه کاربردی دارد؟ آن را به شبکه اضافه کنید و تاثیرش را با درصد های مختلف

برای خطا و دقت بررسی نمایید.

سوال ۲ – Transfer Learning

در این سوال به بررسی معماری های مختلف شبکه عصبی که بر اساس شبکه های CNN بوجود آمده اند را بررسی میکنیم و به پیاده سازی آنها میپردازیم. ابتدا بر اساس رقم آخر شماره دانشجویی خود و یا جمع رقم آخر خود و همگروهیتان مدل را انتخاب نمایید. نمودار های خطا و دقت را برای هر بخش بیاورید.

AlexNet	۰
VGG-۱۶	۱
VGG-۱۹	۲
Inception	۳
GoogLeNet	۴
ResNet	۵
Squeeze Net	۶
DenseNet	۷
Shuffle Net	۸
ENet	۹

الف) معماری شبکه را توضیح دهید و مزایا و معایب آن را بررسی نمایید. همچنین در مورد پیش پردازش های لازم برای داده های ورودی به شبکه بحث کنید.

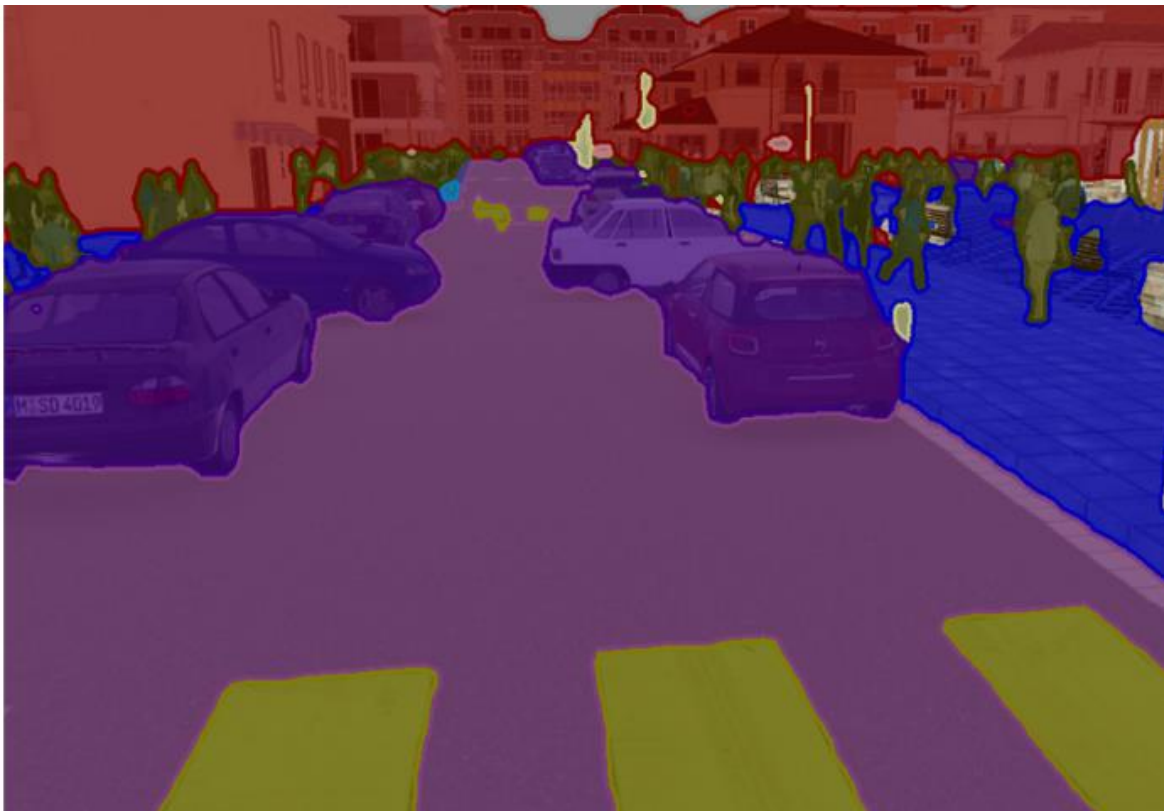
ب) ابتدا شبکه را پیاده سازی کنید. سپس سه عکس بطور دلخواه انتخاب کنید و پس از پیش پردازش های لازم دسته های پیش بینی شده توسط شبکه را مشاهده کنید.

ج) بررسی کنید چه نوع عکس هایی را شبکه شما قابلیت تشخیص آن را دارد. اگر عکسی داخل آن دسته نباشد چه می شود؟ راه حل چیست؟

سوال ۳ – Semantic Segmentation

یکی از کاربرد های شبکه های عصبی در بحث Semantic Segmentation می باشد. هدف از تقسیم بندی تصویر معنایی (Semantic Segmentation) این است که هر پیکسل از یک تصویر را با یک کلاس مربوط به آنچه نشان داده می شود، برچسب گذاری کند. از آنجایی که ما برای هر پیکسل در تصویر پیش بینی می کنیم، این کار معمولاً به عنوان پیش بینی متراکم (dense prediction) شناخته می شود.

از بین دو مدل [UNET](#) و [DEEPLAB](#) یکی را به دلخواه انتخاب کنید و به دو بخش زیر پاسخ دهید



الف) مقاله مربوط به مدل را مطالعه کنید و معماری و کاربرد مدل را بطور کامل توضیح دهید.

ب) مدل را با جست و جو در اینترنت بیابید و بر روی یک عکس بطور دلخواه آزمایش کنید و نتایج را توضیح دهید.

سوال ۴ – Object Detection

در بین مدل های تشخیص شیء، معماری های یک پارچه YOLO از لحاظ سرعت و دقت به نتایج چشم گیری رسیدند. این الگوریتم ها میتوانند در زمان کمتری در مرحله آموزش به همگرایی برسند و در مرحله استنتاج، اشیاء را با سرعت بالایی تشخیص داده و دسته بندی کنند. با استفاده از مطالبی که در ویدئوهای ضبط شده YOLO در اختیار شما قرار گرفته است و نیز منابع موجود در اینترنت به سوالات زیر پاسخ دهید. دقت کنید برای بخش پیاده سازی، شما می توانید از [این نوت بوک](#) استفاده کنید، اما باید مراحل کد را به صورت مختصر شرح دهید.

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

ب) عملکرد YOLOv1 بر روی داده هایی که دارای اشیائی با تداخل (Overlap) بالا هستند، چه مشکلی دارد؟ توضیح دهید. برای حل این مشکل در YOLOv2 چه روشی اتخاذ شده است؟

ج) در مورد پیشرفت های YOLOv4 و YOLOv5 نسبت به ورژن های قبلی توضیح دهید.

د) برای آموزش شبکه ی YOLOv3 بر روی مجموعه داده دلخواه باید بعضی از hyperparameterهای شبکه را متناسب با مجموعه داده تنظیم کنیم. با مشخص کردن این hyperparameterها، نقش هر کدام و نحوه ی مشخص کردن مقدار مناسب آنها را توضیح دهید.

ه) با استفاده از مجموعه داده ای که در اختیار شما قرار داده شده است، یک بار شبکه ی YOLOv5s و یک بار شبکه ی YOLOv5x را آموزش دهید و نتایج هر کدام (شامل $mAP_{0.5}$ ، $mAP_{0.5:0.95}$ ، Recall، Precision و زمان آموزش و تشخیص) را با یکدیگر مقایسه کنید. مزایا و معایب استفاده از مدل بزرگ تر را نسبت به استفاده از مدل کوچک تر را بیان کنید.

نکات:

- مهلت تحویل این مینی پروژه ۱۶ آذر 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 هیچ‌وجه 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 ۲۳ 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.

HW#/PROJECT#_[Lastname]_[StudentNumber].zip

- در صورت وجود هرگونه ابهام یا مشکل می‌توانید از طریق رایانامه‌های زیر با دستیاران آموزشی مربوطه آقایان امیر محمد رادمهر (سوال ۱ تا ۳) و راستین سورکی (سوال ۴) در تماس باشید:

amir.radmehr77@gmail.com

Rustin.soraki@gmail.com