

به نام خدا

دانشکده فنی و مهندسی - گروه مهندسی کامپیوتر

شبکه‌های عصبی مصنوعی - تمرین‌های سری سوم - پاییز 1401

آخرین موعد تحویل 1401/10/15



سوال 1.

(آ) فرض کنید ورودی یک لایه کانولوشنی تصاویری با ابعاد 65×65 و دارای ۳ کانال باشد و لایه کانولوشنی دارای 128 فیلتر کانولوشن 5×5 با گام ۲ باشد. ابعاد خروجی و همچنین تعداد پارامترهای این لایه چقدر است؟
(ب) در قسمت قبل، فرض کنید با یک لایه تمام متصل از تصویر ورودی به خروجی با ابعاد مشابه خروجی قسمت قبل برسیم. تعداد پارامترها را در این حالت نیز به دست آورده و با قسمت قبل مقایسه کنید.
(پ) توضیح دهید چگونه Max-Pooling باعث مقاومت شبکه نسبت به چرخش در تصویر ورودی می‌شود.

سوال 2.

(آ) یک شبکه عصبی برای تشخیص چشم باز و بسته با استفاده از لایه‌های کانولوشنی در مجموعه داده [MRL Eye Dataset](#) آموزش دهید.
(ب) در فرایند آموزش از چه optimizer ی استفاده کردید و چرا؟
(پ) نمودارهای accuracy و loss را برای مرحله آموزش رسم نمایید.
(ت) با استفاده از شبکه آموزش دیده شده برای هر کدام از ویدیوهای موجود در دیتاست [eyeblink8](#) گزارش دهید که چند بار چشم فرد کاملاً بسته بوده است. برای تشخیص مکان چشم می‌توانید از ابزار Dlib یا ابزارهای مشابه استفاده نمایید.
روش به کار برده شده را کامل توضیح دهید.

سوال 3.

مجموعه داده Cifar100 را بارگزاری کنید برای دسترسی به مجموعه داده می‌توانید از کتابخانه `keras.datasets` استفاده نمایید.
(آ) شبکه CNN با فیلترهای 5×5 طراحی کنید و تعداد پارامترهای شبکه را گزارش نمایید. برای توقف آموزش Callback را طوری تنظیم نمایید، تا وقتی مقدار loss در داده‌های validation در شش epoch بهبودی نداشته باشد.
(ب) تابع خطای به کار برده شده را توضیح دهید و دلیل استفاده از آن را بیان نمایید.
(پ) در مرحله بعد سعی کنید با استفاده از Drop out تعداد epoch های شبکه قسمت (آ) را بدون رسیدن به بیش برآزش افزایش دهید. دقت کنید که نتایج این مرحله باید جدا از مرحله قبلی گزارش گردد.
تذکر: نسبت داده‌های آموزش به تست، 5 به 1 قرار گیرد و 20 درصد داده‌های آموزش را برای validation در نظر بگیرید و نمودارهای مربوط به loss و accuracy را برای داده‌های train و validation رسم نمایید. تعداد epoch ها را نیز گزارش نمایید.

سوال 4.



آ) وزن‌های شبکه vgg16 را که توسط مجموعه داده ImageNet آموزش دیده است، در نظر بگیرید. لایه‌های fully connected این شبکه را از آن جدا نمایید و لایه fully connected جدیدی به آن اضافه کنید. داده‌های cifar100 را بر روی این شبکه fine tune نمایید. می‌توانید از کتابخانه keras.datasets استفاده نمایید.

ب) یک بار نیز در شبکه طراحی شده، لایه‌های بلاک پنجم (block5) شبکه vgg16 را برای آموزش دیدن باز بگذارید و نتیجه را گزارش دهید.

پ) دقت کدام روش بیشتر است. توضیح دهید.

تذکر: در تمام مراحل نسبت داده‌های آموزش به تست، 5 به 1 قرار گیرد و 20 درصد داده‌های آموزش را برای validation در نظر بگیرید و نمودارهای مربوط به loss و accuracy را برای داده‌های train و validation رسم نمایید.

سوال 5.

یکی از کاربردهای شبکه‌های عصبی کانولوشنی دسته‌بندی ویدیو می‌باشد. در این تمرین می‌خواهیم از قدرت یک شبکه از پیش آموزش دیده شده برای استخراج ویژگی جهت دسته‌بندی یک مجموعه داده ویدیویی استفاده نماییم. مجموعه داده Hockey dataset را از این [لینک](#) دانلود کنید. شبکه ResNet50 را که از پیش بر روی مجموعه داده ImageNet آموزش دیده‌است را در نظر بگیرید.

آ) با استفاده از شبکه عصبی موجود، یکی از لایه‌ها را برای استخراج ویژگی انتخاب نمایید و فریم‌های هر ویدیو را به عنوان ورودی به شبکه بدهید و ویژگی‌های آن‌ها را به دست آورید.

ب) تا این مرحله ویژگی‌های مکانی مربوط به هر فریم را استخراج کرده‌اید، حال نیاز است تا اطلاعات زمانی در طول فریم‌های متوالی را به دست آورید. فرض کنید یک ماتریس ویژگی برای هر ویدیو با ابعاد $M \times N$ دارید که در آن M نشان‌دهنده‌ی تعداد ویژگی‌های استخراج شده برای هر فریم است. مقدار M با توجه به لایه‌ای که برای استخراج ویژگی انتخاب کرده‌اید مشخص خواهد شد. همچنین N تعداد فریم‌های مربوط به هر ویدیو می‌باشد. حال با استفاده از عمل میانگین‌گیری در بین ویژگی‌های هر فریم به دست آمده با ایندکس یکسان یک ماتریس منحصر به فرد با سایز $M \times 1$ به دست آورید.

پ) یک شبکه تمام متصل برای مرحله کلاس‌بندی طراحی کنید با استفاده از نیمی از داده‌های هر دسته فرایند train را انجام دهید و با داده‌های باقی‌مانده، دقت روش خود را بر روی داده‌های تست ارزیابی کنید.

ت) کدام لایه شبکه را برای استخراج ویژگی انتخاب کرده‌اید و چرا؟ تعداد ویژگی‌های استخراج شده برای هر فریم چقدر است؟ توجه داشته باشید که شرح تمامی مراحل در گزارش نوشته‌شده ضروری می‌باشد.

برای درک بهتر ویژگی‌های زمانی و مکانی در ویدیو بررسی مقاله "[Pooled Motion Features for First-Person Videos](#)" می‌تواند مفید باشد.

توضیحات:



- ✓ پاسخ تمرین‌ها را در قالب یک فایل فشرده با نام `firstname_lastname_studentnumber_HW3` به آدرس ایمیل khu.ann.homework@gmail.com با عنوان EXC3 ارسال نمایید.
 - ✓ اگر به مشکل پیش‌بینی نشده‌ای در حل مسائل روبه‌رو شدید، مشکل را در گزارش خود مطرح کنید و راه حل را بیان کنید.
 - ✓ همکاری، همفکری و کمک گرفتن در خصوص چگونگی انجام تمرین مانعی ندارد، اما تهیه کدها و گزارشات باید به صورت انفرادی انجام شود (تشابه گزارش‌ها و کدهای ارسالی، نشان‌دهنده عدم توجه به این تذکر است).
 - ✓ توجه داشته باشید که تمامی موارد خواسته شده در سوالات را تا حد امکان در گزارش خود توضیح دهید تا حقی از شما ضایع نگردد.
 - ✓ دقت داشته باشید که برای گرفتن نمره کامل باید در موعد معین شده پاسخ‌های خود را ارسال کنید. به ازای هر روز تاخیر 10 درصد نمره کسر خواهد شد و پس از چهار روز نمره‌ای برای شما در نظر گرفته نخواهد شد. (آخرین مهلت ارسال ساعت 24 روز 20 دی می باشد و بعد از این زمان هیچ نمره‌ای منظور نخواهد شد)
 - ✓ تمرین‌های پیاده‌سازی را در قالب فایل نوتبوک پایتون (.ipynb) ارسال کنید و نتیجه اجرا هر سلول باید در زیر آن قرار گرفته باشد. همچنین نوشتن گزارش، شامل موارد خواسته شده و نمودارهای به دست آمده ضروری می‌باشد.
- موفق باشید