

اصول بینایی کامپیوتر (۱- ۲۲۹۲۶) زمستان ۱۳۹۹
تمرینات سری پنجم
مهلت تحویل: ۲۶ تیر ۱۴۰۰ (شنبه) ساعت ۱۲ شب

لطفاً به نکات زیر توجه بفرمایید.

۱. نتایج و پاسخ های خود را در یک فایل zip (rar نباشد) در سایت cw قرار دهید (ایمیل نکنید). در صورت رعایت نکردن این موارد، ۱۰ نمره از شما کسر خواهد شد.

۲. اغلب تمرینات نیاز به برنامه نویسی خواهند داشت. در چنین تمرین هایی، کسب نمره کامل در هر سؤال مستلزم تحویل سه مورد نتایج، کدها، و توضیحات می باشد. نتایج مورد نیاز در هر تمرین توضیح داده شده است. نتایج ۳۰ درصد نمره شما را تشکیل می دهند. حتی با وجود توضیحات کامل و کدهای قابل اجرا، اگر نتایج در بین فایل های شما نباشند این ۳۰ درصد به شما تعلق نمی گیرد. در مورد کدها و توضیحات در زیر توضیح بیشتری داده میشود.

۳. برای سؤالاتی که نیاز به برنامه نویسی دارند، باید حتماً کدهای استفاده شده که منجر به نتایج فرستاده شده است همراه فایل های شما باشند. با اجرای این کدها باید همان نتایجی که فرستاده اید قابل بازیابی باشند. برنامه اصلی شما باید با نام مرتبط با شماره سؤال ذخیره شده باشد که در هر سؤال به آن اشاره خواهد شد. برنامه شما باید به گونه ای باشد که بدون نیاز به هیچ تغییری قابل اجرا باشد، در غیر اینصورت هیچ نمره ای تعلق نخواهد گرفت. کدهای شما ۴۰ درصد نمره هر سؤال را تشکیل می دهند و در صورت عدم وجود و یا کار نکردن کد این نمره به شما تعلق نخواهد گرفت. در صورت استفاده از فایل های متعدد لطفاً تمام آن ها را به همراه پاسخ های خود بفرستید تا برنامه شما قابل اجرا باشد. در چنین مواردی، می توانید فایل ها را با نام های دلخواه خود ذخیره نمایید ولی فایل اصلی باید با نام اشاره شده در هر سؤال ذخیره شود و طوری باشد که با اجرای آن برنامه تمام قسمت های برنامه مورد نظر اجرا شود. در صورتی که چند کد در یک سؤال از شما خواسته شده باشد، باید تمام آن ها را با توضیحات خواسته شده در سؤال ذخیره نموده و بفرستید. کدهای شما تماماً باید توسط خودتان نوشته شده باشند. هرگونه استفاده از کد دیگران، اعم از دوستان و اینترنت، به هر شکل ممکن، اعم از کپی کردن یا همکاری کردن، تقلب محسوب می شود و نمره تمام تمرینات جاری و تمام تمرینات قبلی صفر خواهد شد.

۴. برای تمام سؤالات، باید جزئیات روشی که استفاده کرده اید را توضیح دهید. این توضیحات برای تمام سؤالات می توانند در یک فایل pdf باشند. این قسمت ۳۰ درصد نمره هر سؤال شما را تشکیل می دهد. در توضیحات، باید اشاره کامل به کارهایی که انجام داده اید بنمایید به طوری که یک شخص آگاه از موارد درس بتواند به آسانی متوجه کاری که شما انجام داده اید بشود.

۵. تمام فایل های مربوط به یک سری تمرین را باید با هم تحویل دهید. در صورتیکه قسمت های مختلف یک سری تمرینات را در زمان های مختلف در سایت cw قرار داده باشید، آخرین زمان بارگزاری به عنوان تاریخ تحویل شما در نظر گرفته خواهد شد.

سؤالات:

تشخیص صحنه با استفاده از شبکه عصبی (Scene Recognition Using Neural Networks) (۱۰۰ نمره)

سوال ۴ در تمرین سری ۳ را این بار با استفاده از شبکه های عصبی حل کنید. از همان داده ها با همان داده آموزش و تست استفاده نمایید. داده تست در واقع داده اعتبار سنجی برای به دست آوردن بهترین مدل و پارامترها در این مساله می باشد ولی از آنجاییکه در این مساله داده تست نداریم از اصطلاح تست به جای اعتبارسنجی استفاده می کنیم. برای این تمرین توصیه می شود از کتابخانه های Tensorflow یا Pytorch استفاده کنید. منابع بسیاری در اینترنت برای راهنمایی برای استفاده این دو کتابخانه می توانید پیدا کنید.

در این تمرین از معماری شبکه AlexNet که روی داده ImageNet برای مساله دسته بندی (classification) آموزش دیده است در حالت های مختلف استفاده خواهید نمود. شبکه AlexNet برای مساله دسته بندی به ۱۰۰۰ کلاس آموزش دیده است، در نتیجه خروجی لایه آخر آن یک بردار ۱۰۰۰ تایی می باشد. ورودی لایه آخر که همان خروجی لایه هشتم می شود یک بردار ۴۰۹۶ تایی می باشد. این بردار را می توان در واقع به عنوان بردار ویژگی های تصویر ورودی در نظر گرفت. برای استخراج بردار ویژگی های یک تصویر و یا یک زیر تصویر روش های مختلفی طراحی شده اند که برخی از آن ها در کلاس مطرح شدند. طراحی یک روش خوب برای استخراج بردار ویژگی های با معنی و خوب کار بسیار چالش برانگیزی است. شبکه AlexNet و شبکه های دیگر مشابه در واقع عمل استخراج بردار ویژگی ها را خودشان انجام می دهند بدون اینکه ما به صورتی دستی روش خاصی برای این کار تعریف کرده باشیم. از آنجاییکه تعریف های ما برای استخراج بردار ویژگی همیشه ممکن است دارای نقص هایی باشند و برخی موارد در نظر گرفته نشده باشند، چنانچه داده های بسیار زیاد و متنوعی به شبکه هایی مانند AlexNet نشان داده شود چه بسا بردار ویژگی که به دست می آورند بهتر از آنهایی باشد که متخصصین تعریف می کنند. بنابراین، خروجی لایه هشتم که یک بردار ۴۰۹۶ تایی است را می توانیم به عنوان بردار ویژگی تصویر ورودی در نظر بگیریم. در این صورت، لایه آخر که خروجی آن یک بردار ۱۰۰۰ تایی است در واقع یک روش دسته بندی روی ۱۰۰۰ دسته با استفاده از بردارهای ویژگی خروجی لایه هشتم می باشد. در مسائل مختلف دیده شده است که این بردار ویژگی ها حتی برای مسائلی که دسته بندی نیستند و یا دسته ها متفاوت از دسته های ImageNet هستند هم با معنی بوده و به نتایج خوب منجر می شوند.

برای استفاده از این بردارهای ویژگی در یک مساله دسته بندی دیگر، کافی است خروجی لایه آخر را برداشته و به جای آن یک بردار خروجی به اندازه تعداد دسته ها در مساله جدید قرار دهیم. در این صورت، وزن های ۸ لایه اول می توانند همان وزن های AlexNet که روی ImageNet آموزش دیده شده است باشند ولی وزن های لایه آخر را باید به دست آوریم. می توان وزن های ۸ لایه اول را بدون تغییر نگاه داشت و با استفاده از داده آموزش جدید در مساله جدید شبکه را آموزش دهیم به طوری که تنها وزن های لایه آخر به روز رسانی شوند. در این صورت، در واقع از ۸ لایه اول استفاده می شود تا بردارهای ویژگی تصویر ورودی به دست آید، سپس لایه آخر مساله دسته بندی را حل می کند.

کار دیگری که معمولاً منجر به جواب های بهتری می شود این است که در شبکه جدید وزن های ۸ لایه اول را از AlexNet آموزش داده شده روی ImageNet مقداردهی اولیه نماییم و لایه آخر را به صورت رندوم مقداردهی اولیه نماییم، و سپس تمام وزن های شبکه را با داده آموزش مساله جدید به روز رسانی کنیم. در این صورت در واقع مقدار اولیه خوب برای مقدار وزن ها در نظر گرفته شده است و انتظار داریم که مقدار بهینه وزن ها برای مساله جدید نزدیک مقدار وزن های AlexNet باشند. به همین دلیل در این روش از مقدار سرعت یادگیری (learning rate) بسیار کم استفاده می شود. به این کار که وزن

های شبکه را از مساله دیگری که در آنجا شبکه خوب کار کرده است به یک مساله مشابه دیگر منتقل کنیم و در مساله جدید از آنها استفاده کنیم یادگیری انتقال (transfer learning) گفته می شود. به این کار که وزن های شبکه انتقال یافته را کمی به روز رسانی کنیم تا مقادیر بهینه برای مساله جدید به دست آید تنظیم دقیق (fine tuning) گفته می شود.

اگر داده آموزش مساله جدید کم باشد قادر به آموزش دادن یک شبکه بزرگ مانند AlexNet با مقدار وزن های اولیه رندوم نخواهیم بود. در این تمرین تعداد داده آموزش بسیار کم است و برای آموزش دادن شبکه ای به بزرگی AlexNet کافی نیست. بنابراین، اگر معماری شبکه AlexNet را در نظر بگیرید و وزن های آن را به صورت رندوم مقدار دهی اولیه کنید و سپس با داده آموزش این تمرین آموزش دهید به جواب خوبی نخواهید رسید.

اگر شبکه در نظر گرفته شده بسیار ساده باشد، شبکه قادر نخواهد بود که این مساله را به خوبی حل کند. بنابراین لازم است که از شبکه های به اندازه کافی پیچیده مانند AlexNet برای این مساله استفاده شود. از آنجاییکه تعداد داده آموزش کافی نیست، لازم است که از انتقال یادگیری و تنظیم دقیق پارامترها استفاده شود.

در این تمرین، مواردی که در بالا توضیح داده شد را تجربه خواهید کرد. مراحل زیر را انجام دهید. ابتدا تمام توضیحات را تا آخر بخوانید.

۱- در ابتدا یک شبکه ساده در نظر می گیرید. معماری شبکه AlexNet را در نظر بگیرید. از یک لایه کانولوشن اول و دو لایه اتصال کامل (fully connected) آخر استفاده نموده و یک شبکه جدید تشکیل دهید. وزن های این شبکه را به صورت رندوم مقداردهی اولیه نموده و آموزش دهید. در لایه کانولوشن در قسمت تجمیع حداکثری (max pooling) از سلول های به اندازه ۴ با اندازه گام (stride) به بزرگی ۴ استفاده نمایید تا خروجی آن کوچک تر شود.

۲- در معماری شبکه AlexNet سه لایه کانولوشن اول و سه لایه اتصال کامل (fully connected) آخر را در نظر گرفته و یک شبکه جدید تشکیل دهید. وزن های این شبکه را به صورت رندوم مقداردهی اولیه نموده و آموزش دهید. در لایه کانولوشن سوم به جای ۳۸۴ فیلتر از ۲۵۶ فیلتر استفاده کنید. همچنین به این لایه تجمیع حداکثری با اندازه سلول ۲ و گام ۲ اضافه نمایید تا اندازه خروجی مشابه خروجی لایه پنجم AlexNet شود.

۳- معماری کامل شبکه AlexNet را در نظر بگیرید. وزن های این شبکه را مقداردهی اولیه نمایید و شبکه را آموزش دهید.

۴- معماری کامل شبکه AlexNet را در نظر بگیرید. وزن های ۸ لایه اول را از شبکه AlexNet اصلی که روی داده ImageNet برای مساله دسته بندی (classification) آموزش دیده است کپی کنید و وزن های لایه آخر که به منظور انجام دسته بندی در مساله جدید است را با مقادیر رندوم مقداردهی اولیه نمایید. حال شبکه را آموزش دهید به این صورت که وزن ۸ لایه اول تغییر نکند و فقط وزن های لایه آخر تغییر کنند.

۵- همان کار قسمت قبل را انجام دهید با این تفاوت که وزن های ۸ لایه اول هم تنظیم دقیق شوند.

نتایج به دست آمده در پنج حالت بالا را در گزارش خود ذکر کرده و تحلیل نمایید. در هریک از مراحل می توانید به صلاح دید خود کارهایی که فکر می کنید می تواند باعث نتیجه بهتر شوند را امتحان کنید و در گزارش خود ذکر نمایید. برای مثال

می توانید از حذف تصادفی (dropout) یا نرمال کردن داده ها در خروجی لایه ها یا نرمال کردن داده ورودی به شبکه استفاده نمایید. برای هر یک از پنج حالت بالا کارهای زیر را انجام دهید.

سعی کنید داده آموزش را افزایش دهید. می توانید هر تصویر را را راستای افقی آینه کنید، تصاویر جدید با اعمال مقدار کمی پرسپکتیو یا دوران به دست آورید، و یا از هر روش دیگری که فکر می کنید می تواند در این مساله موثر باشد استفاده کنید. کارهای انجام شده را در گزارش خود توضیح دهید.

برای آموزش دادن شبکه های خود، هر شبکه را حداقل ۳۰ اپاک (epoch) آموزش دهید. یک اپاک به معنای این است که در فرآیند آموزش شبکه تمام داده آموزش را دقیقاً یک بار دیده است. در هر اپاک مقدار خسارت (loss) روی داده آموزش و تست را به دست آورید و روی یک نمودار نمایش دهید. محور افقی نمودار شماره اپاک و محور عمودی مقدار خسارت را نشان می دهند. در ابتدا با وزن های اولیه و سپس در انتهای هر اپاک مقدار خسارت داده آموزش را به دست آورده و روی نمودار نشان دهید. این نقاط را به هم وصل کنید تا یک نمودار برای خسارت داده آموزش در طول زمان با جلو رفتن آموزش شبکه به دست آید. یک نمودار مشابه هم برای داده تست در همان نمودار بکشید. خسارت داده تست، به استثناء ابتدای آموزش، همواره باید بیشتر از داده آموزش باشد. اگر آموزش شما به خوبی پیش برود، خسارت آموزش و تست هر دو رفته رفته کم می شوند، گرچه خسارت تست همواره بیشتر از آموزش باقی می ماند. در صورتی که مدل به داده آموزش بیش برآزیده (overfit) شود، خسارت تست از یک جایی به بعد بیشتر می شود در حالیکه خسارت آموزش کمتر شده و یا حتی ممکن به صفر نزدیک شود. بهترین پارامترها برای این شبکه شما جایی به دست می آید که خسارت تست کمینه شده است و بعد از آن افزایش پیدا می کند. در صورت چنین اتفاقی سعی کنید با روش هایی که بلد هستید، مانند حذف تصادفی، از بیش برآزیده شدن جلوگیری کنید.

نمودارهایی همانند نمودار خسارت برای دقت شبکه ها هم بکشید. در هر اپاک دقت شبکه را به دست آورده و مانند نمودار خسارت نمودار دقت را نیز بکشید. نمودار دقت را برای دو حالت بکشید: یکی دقت در بهترین تشخیص (top1)، و دیگری دقت در ۵ بهترین تشخیص (top5). در اولین حالت تشخیص شبکه را درست در نظر بگیرید اگر دسته با بیشتری مقدار خروجی در لایه خروجی درست باشد. در حالت دوم تشخیص شبکه را درست در نظر بگیرید اگر دسته درست بین ۵ دسته ای که بیشترین امتیاز را در بردار خروجی آورده اند باشد.

کارهای انجام شده را به طور کامل در گزارش خود توضیح دهید. بهترین دقت به دست آمده برای هر مرحله را در گزارش خود ذکر کنید. نمره شما به میزان دقت شبکه های شما به طور کلی و در مقایسه با بقیه دانشجویان کلاس بستگی خواهد داشت. کد اصلی خود را با نام HW5 ذخیره نمایید.