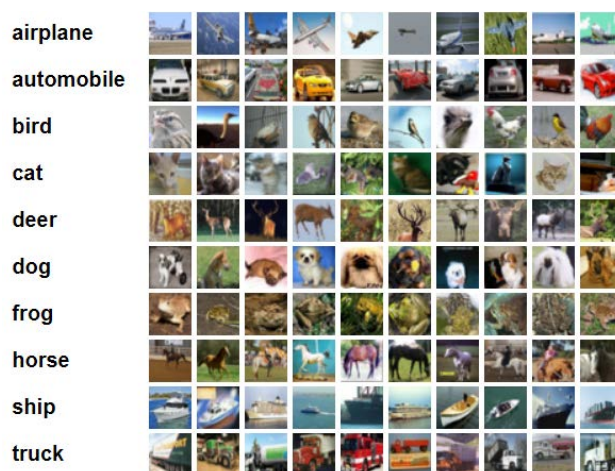


۱. (۱۰٪) [پژوهش] تاکنون بسترهای مختلفی مانند TensorFlow، Torch و Keras با زبانهای برنامه‌نویسی مختلف برای پیاده‌سازی الگوریتم‌های یادگیری عمیق ارائه شده است. گزارشی از مطالعه این بسترها و مقایسه آنها با همدیگر را ارائه دهید و مشخص کنید هر کدام از آنها کدامیک از الگوریتم‌های یادگیری ماشین را دارا هستند.

۲. (۹۰٪) [پیاده‌سازی: بازشناسی اعداد با یادگیری عمیق] در این سوال قرار است که با استفاده از یک چارچوب برنامه‌نویسی یادگیری عمیق، یک مدل برای بازشناسی اعداد صفر تا نه از روی تصاویر اسپکتروگرام که دادگان آن در تمرین‌های قبلی در اختیار شما قرار گرفته است، ایجاد کنید. برای این کار ابتدا یک شبکه برای دسته‌بندی تصاویر (غیر اعداد) ایجاد می‌کنیم که در آن از مجموعه داده‌های CIFAR-10 استفاده کنیم. این مجموعه داده از ۶۰ هزار تصویر رنگی با ابعاد ۳۲×۳۲ تشکیل شده است (شامل ۵۰ هزار تصویر آموزشی و ۱۰ هزار تصویر آزمایشی). تصاویر این مجموعه داده در ده دسته قرار گرفته‌اند که نمونه‌ای از هریک از این دسته‌ها در شکل ۱ ملاحظه می‌شود.



شکل ۱ نمونه تصاویری از هریک از برجسب‌های مجموعه داده‌ی CIFAR-10



به صفحه‌ی رسمی این مجموعه داده^۱ مراجعه کنید، آن را دریافت کرده و مجموعه‌های آزمایشی و آموزشی آن را در این بخش از تمرین مورد استفاده قرار دهید. داده‌های آموزشی این پایگاه داده در پنج دسته‌ی جداگانه (هریک با اندازه‌ی ۱۰ هزار تصویر) ارائه شده‌اند. در صورتی که منابع شما اجازه‌ی اجرای الگوریتم بر روی تمامی مجموعه داده را نمی‌دهد، می‌توانید به عنوان داده‌ی آموزشی، تنها دسته اول آن (data_batch_1) را مورد استفاده قرار دهید. توجه کنید که در این صورت لازم است این موضوع را در گزارش خود ذکر کنید. شبکه‌ای با معماری زیر در نظر بگیرید:

- (۱) لایه‌ی پیش‌بینی: اندازه‌ی کرنل (فیلتر): 3×3 ، (تعداد کانال‌های خروجی: ۷ و تابع فعال‌سازی ReLU)
- (۲) لایه‌ی پیش‌بینی: اندازه‌ی کرنل: 3×3 ، (تعداد کانال‌های خروجی: ۹ و تابع فعال‌سازی ReLU)
- (۳) لایه‌ی ادغام بیشینه: اندازه‌ی کرنل: 2×2
- (۴) لایه drop-out با احتمال ۲۵ درصد
- (۵) لایه خطی با تعداد ده (تعداد دسته‌ها) نرون خروجی

این شبکه را با بهینه‌ساز Adam، تابع خطای categorical cross entropy و نرخ یادگیری یک صدم آموزش دهید (با اندازه دسته ۳۲). این معماری و نحوه‌ی آموزش را به عنوان حالت پایه در نظر گرفته و در هریک از بخش‌های الف تا خ تغییرات گفته شده را بر روی همین معماری و پارامترها اعمال کنید. قسمتی از داده‌های آموزشی را به عنوان داده‌های validation جدا کرده (می‌توان به عنوان یک پارامتر هنگام فراخوانی تابع آموزش نیز تنظیم شود) و میزان خطای آن را معیاری برای زمان توقف الگوریتم قرار دهید. همچنین از دقت داده‌های آزمایشی به منظور ارزیابی مدل استفاده کنید.

الف) دقت مدل را روی داده آزمایش CIFAR-10 گزارش کنید. همچنین، نموداری از روند تغییرات خطای شبکه در حین آموزش، بر روی داده‌های train و validation (این دو نمودار را روی یکدیگر قرار دهید) و میانگین مدت زمان اجرای تکرار^۲ها را گزارش کنید.

ب) حال برای همه سیگنال‌های گفتار مربوط به اعداد دادگان تمرین‌های قبلی، اسپکتروگرام را تولید کنید (تولید هر سیگنال به یک تصویر) و با شروع از مدل آموزش یافته قسمت الف به عنوان مدل پیش‌آموزش^۳

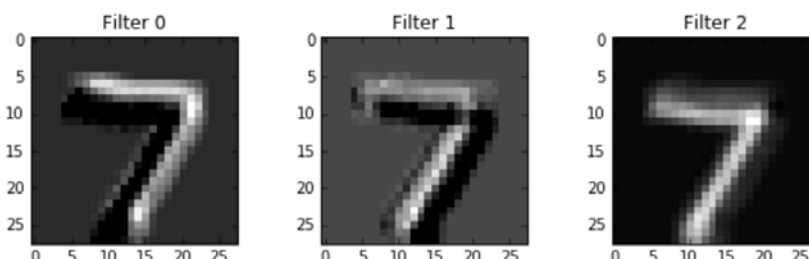
^۱ <https://www.cs.toronto.edu/~kriz/cifar.html>

^۲ epoch

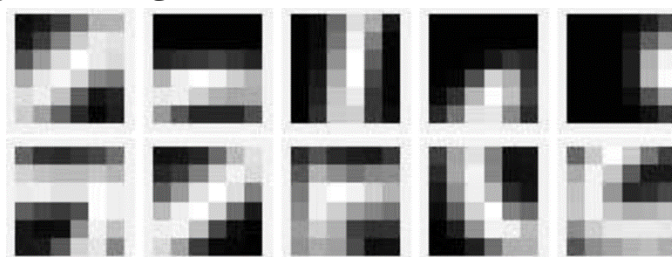
^۳ PreTrain

شبکه را برای تشخیص اعداد بازآموزش دهید. برای این کار، لایه خروجی شبکه را معادل اعداد صفر تا نه قرار دهید و به جز وزن‌های لایه خروجی، بقیه وزن‌های اولیه شبکه را برابر با وزن‌های بدست آمده در بخش الف قرار دهید. شبکه را بخش آموزش داده اعداد آموزش دهید و پس از آن روی داده آزمایش اعداد، دقت، نمودارهای تغییرات خطای درحین آموزش بر روی train و validation و مدت زمان اجرای تکرارها را گزارش کنید.

پ) پس از آموزش مدل، یک تصویر آزمایشی (از دادگان CIFAR-10 و نه از اسپکتروگرام‌ها) را در نظر بگیرید، آن را نمایش دهید و آن را به عنوان ورودی به شبکه بدهید. پس از عبور تصویر از لایه‌ی اول پیچشی، هریک از کانال‌ها تصویری خاکستری^۴ با ویژگی خاصی از تصویر ورودی را نمایش می‌دهد (شبیه به شکل ۲). همچنین وزن‌های مربوط به یکی از کانال‌های کرنل را نیز می‌توان به عنوان یک تصویر خاکستری در نظر گرفت و آن را نمایش داد (شبیه به شکل ۳). حال شما تصویر کانال‌های مختلف لایه اول پیچشی را در کنار تصویر وزن‌های کرنل‌های همان کانال نمایش دهید و ارتباط آن‌ها را توضیح دهید.



شکل ۲ نمونه سه تصویر که توسط سه فیلتر مختلف در لایه‌های میانی شبکه‌ی عصبی تولید شده است



شکل ۳ نمونه ۱۰ تصویر از وزن‌های کرنل (فیلتر) با اندازه‌ی ۶×۶

ت) حالا می‌توانید پارامترهای مختلف شبکه را برای بهینه کردن آن تغییر دهید. از جمله پارامترهای قابل تغییر شبکه تعداد لایه‌ها، اندازه‌ی کرنل، تعداد فیلتر، اندازه گام، نرخ یادگیری، ترتیب چیدن لایه‌ها، تابع

^۴ gray-scale

بِسْمِ اللَّهِ

پروازش گفتار (۸۳-۰۵-۱۴۸)
نیمسال دوم ۱۳۹۸-۱۳۹۹

تاریخ تحویل: ۱۳۹۹/۰۳/۳۰

تمرین شماره ۵



فعال سازی، بهینه سازی، اندازه ی $batch-size$ و ... است. در این قسمت، مراحل الف و ب را تکرار کنید و در آن اندازه ی کرنل لایه اول پیچشی را به 5×5 تغییر دهید، لایه دوم پیچشی را حذف کنید و نرخ یادگیری را به 10^{-4} تغییر دهید. روی داده آزمایش اعداد، دقت، نمودارهای تغییرات خطای درحین آموزش بر روی $train$ و $validation$ و مدت زمان اجرای تکرارها را گزارش کنید.

راهنمایی: برای پیاده سازی ها می توانید از چارچوب های برنامه نویسی $tensorflow$ و $pytorch$ استفاده کنید و برای انجام اجراها که نیازمند حافظه و پردازنده مناسب هستند، از $Google Colab$ بهره بگیرید.