



گزارش تمرین پنجم درس شبکه‌های عصبی



فاطمه غلامزاده

۹۹۱۳۱۰۰۳

سوال اول

برای انتقال یادگیری در شبکه‌های عصبی ۳ روش وجود دارد:

۱. به طور کامل از شبکه پیش آموزش داده استفاده کنیم به گونه‌ای که ورودی را به آن می‌دهیم و خروجی را مستقیماً به عنوان خروجی مسئله خودمان تلقی می‌کنیم. استفاده از این روش فقط در صورتی ممکن است که خروجی‌های شبکه از پیش آموزش داده شده با خروجی‌های مسئله مورد نظر ما یکسان باشد.
۲. در حالتی که خروجی شبکه از پیش آموزش دیده با خروجی مدنظر مسئله ما متفاوت است نیاز است تا خودمان یک classifier جدید را آموزش داده و جایگزین کنیم. در این حالت در واقع از شبکه پیش آموزش دیده به عنوان یک استخراج کننده ویژگی استفاده می‌کنیم و وزن‌های بخش feature extractor در شبکه از پیش آموزش دیده بدون تغییر می‌مانند.
۳. در حالت سوم علاوه بر دسته‌بند، استخراج کننده ویژگی را نیز مختص مسئله خودمان آموزش می‌دهیم که به این عمل fine-tuning یا تنظیم دقیق وزن‌ها گفته می‌شود. در این حالت ابتدا قسمت دسته‌بند را با وزن‌های ثابت در قسمت استخراج کننده ویژگی آموزش می‌دهیم و بعد مجدداً برخی از لایه‌های قسمت feature extractor را از حالت فریز در آورده و مجدداً با نرخ یادگیری بسیار پایین، آن‌ها را آموزش می‌دهیم.

در fine-tuning علاوه بر آموزش دسته‌بند (classifier)، لایه‌های بالایی مدل از پیش آموزش دیده را هم آموزش می‌دهیم. این آموزش موجب می‌شود که وزن‌ها از feature map‌های عمومی و کلی، به فیچرهای متناسب با مجموعه داده مورد استفاده وفق پیدا کنند.

علت fine-tune یا تنظیم دقیق وزن‌ها:

- افزایش سرعت آموزش
- غلبه بر اندازه کوچک دیتاست

- عملکرد مدل را بالا ببریم و دقت مدل افزایش پیدا کند

بنابراین زمانی که دیتاست کوچکی داریم یا اینکه زمان آموزش مدل بالاست استفاده از fine-tuning می‌تواند مناسب باشد. البته داده‌هایی که مدل از پیش آموزش دیده بر روی آن آموزش داده شده است نباید با داده‌های مسئله ما خیلی متفاوت باشد. اما به عنوان مثال اگر دیتاست ما شامل عکس‌های خاصی مانند تصاویر پزشکی یا دست خط چینی باشد و هیچ شبکه از پیش آموزش دیده‌ای در این حوزه‌ها وجود نداشته باشد باید شبکه را از صفر آموزش بدهیم و نمیتوانیم fine-tuning داشته باشیم.

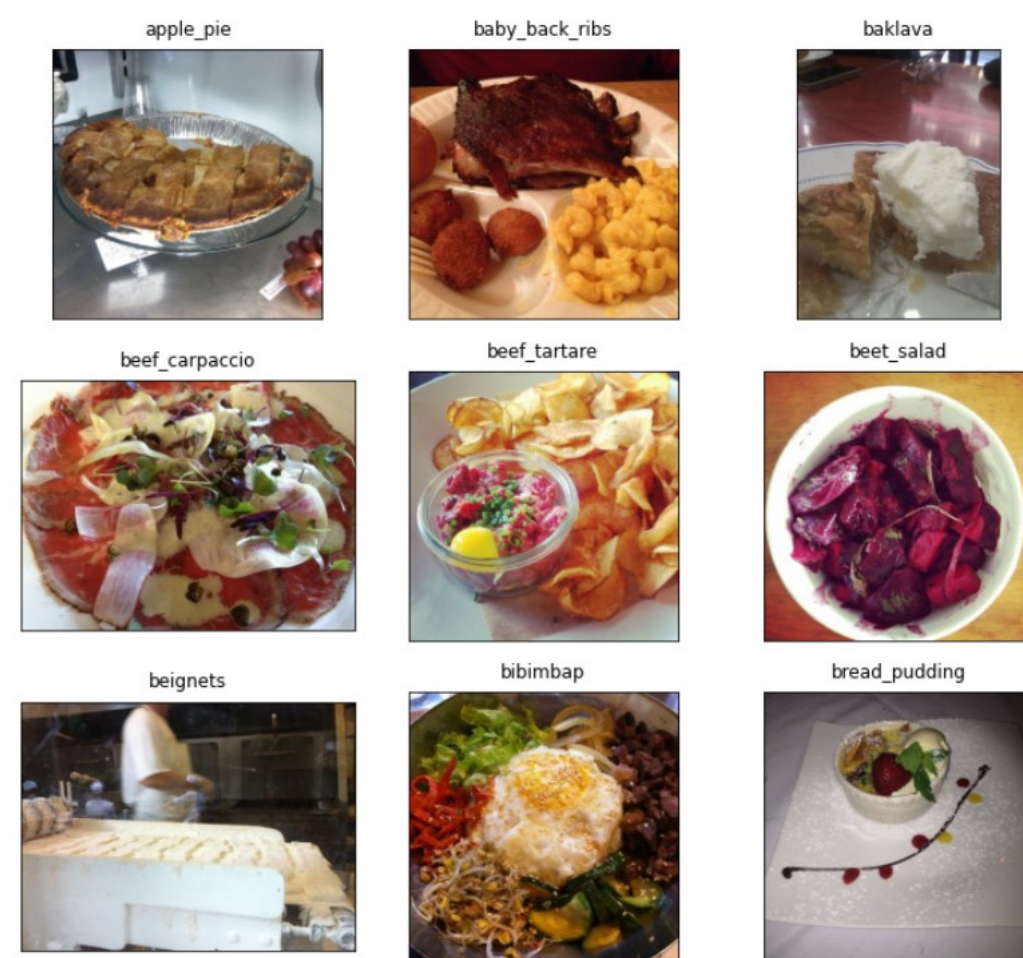
سوال دوم

از آنجایی که لایه‌های ابتدایی در یک شبکه کانولوشنی ویژگی‌ها و فیچرهای کلی و اولیه مانند خطوط عمودی و افقی و لبه‌های تصویر را مشخص می‌کنند و لایه‌های انتهایی هستند که فیچرهای جزئی و مختص مجموعه داده را استخراج می‌کنند، بهتر است برای tuning تنها لایه‌های انتهایی را مجدداً آموزش دهیم چون عملاً آموزش لایه‌های ابتدایی کارایی خاصی نخواهد داشت.

به عنوان مثال می‌توانیم ۱۰ تا ۲۰ درصد از لایه‌های آخر قسمت استخراج کننده ویژگی‌ها را در نظر بگیریم و با آزمون و خطا با گام‌های چندتایی تعداد لایه‌های مختلف را تست کنیم و مدل از پیش آموزش دیده شده را برای مسئله خود اختصاصی کنیم. نرخ یادگیری در tuning بسیار مهم است و این عمل باید با learning rate بسیار کمی (حدوداً یک دهم نرخ یادگیری در حالت عادی) انجام شود.

سوال سوم

چند نمونه از تصاویر به همراه برچسب آن‌ها:

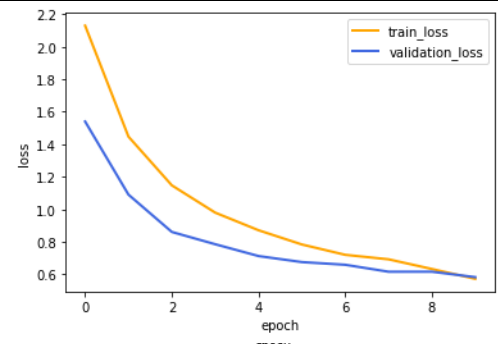
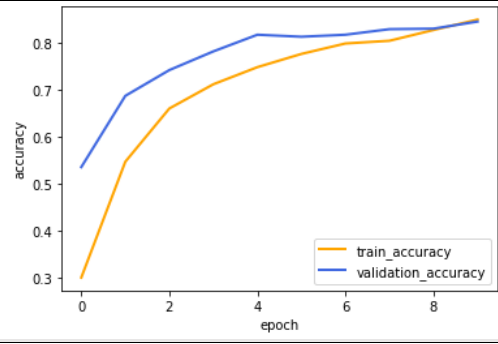
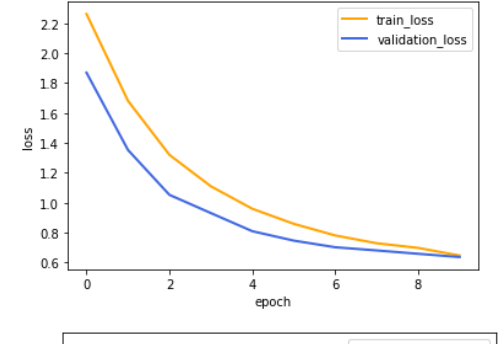
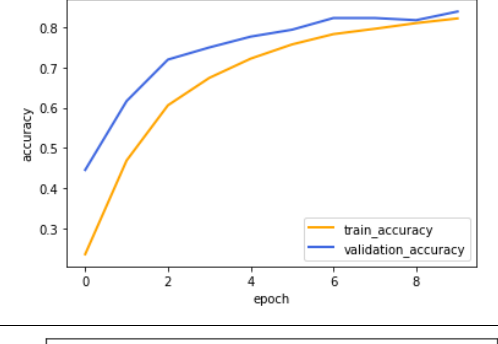
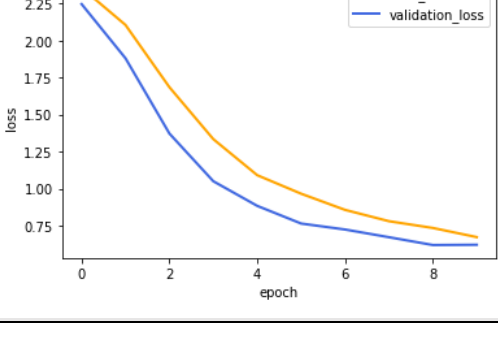
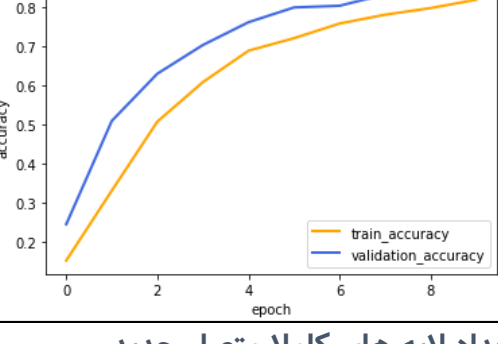


سوال چهارم

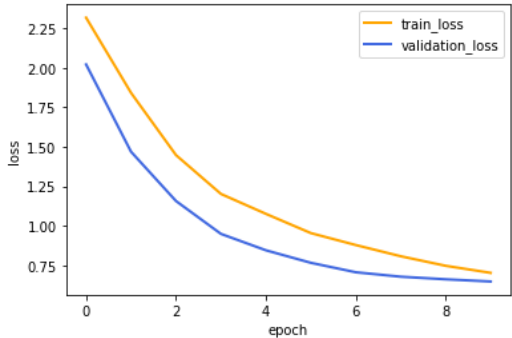
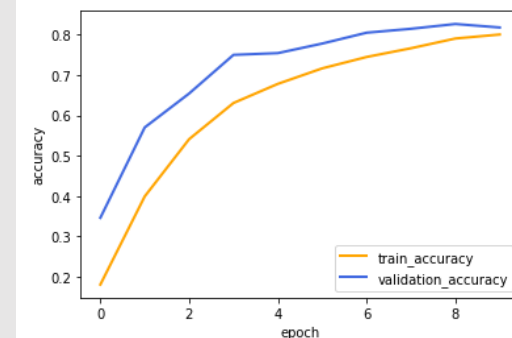
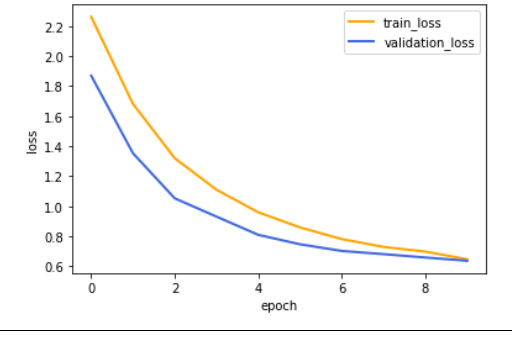
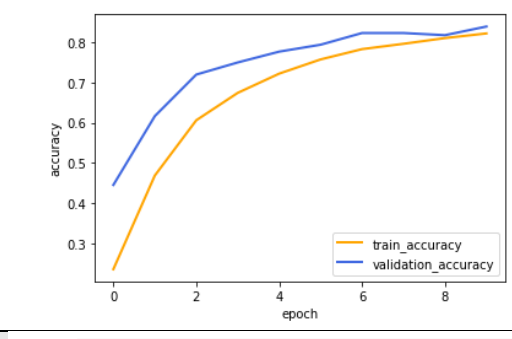
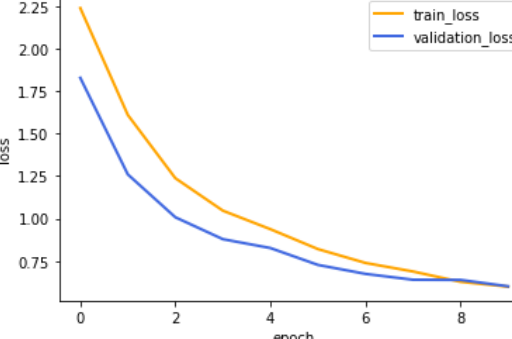
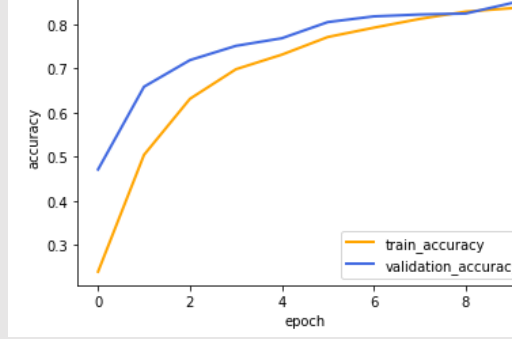
برای انجام این بخش از شبکه Densenet121 استفاده شده است. این شبکه دارای ۴۲۷ لایه می باشد.

بخش اول: بررسی تعداد لایه های کاملاً متصل جدید

در جدول شماره ۱ تعداد لایه های کاملاً متصل جدید با ۳ عدد ۲ و ۳ و ۵ بررسی شده است. همانطور که مشاهده می شود ۳ لایه دارای بیشترین دقت است. افزایش تعداد لایه ها ابتدا باعث افزایش دقت و سپس موجب کاهش دقت شده است. یک توجیه می تواند این باشد که ابتدا مدل ساده است و **underfit** می شود و بعد وقتی تعداد لایه ها زیاد می شود پیچیده می شود و **overfit** می شود و یک حد میانه بیشترین دقت را دارد.

نمودار خطا	نمودار دقت	خطا	دقت	تعداد نورون‌های هر لایه	تعداد لایه‌ها
		0.4823	0.8692	64 10	2
		0.4791	0.8816	128 64 10	3
		0.5131	0.8636	1024 512 128 64 10	5

جدول ۱ بررسی تعداد لایه های کاملاً متصل جدید

نمودار خطا	نمودار دقت	خطا	دقت	تعداد نورون‌های هر لایه	تعداد لایه‌ها
		0.4872	0.8692	64 32 10	3
		0.4791	0.8816	128 64 10	3
		0.4952	0.8688	256 128 10	3

جدول ۲ بررسی تعداد نورون‌ها

ادامه بخش اول: بررسی تعداد نورون‌ها

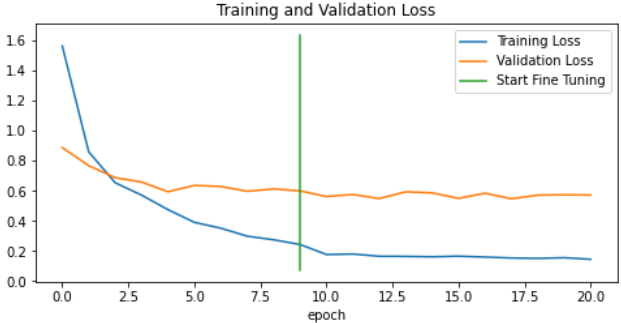
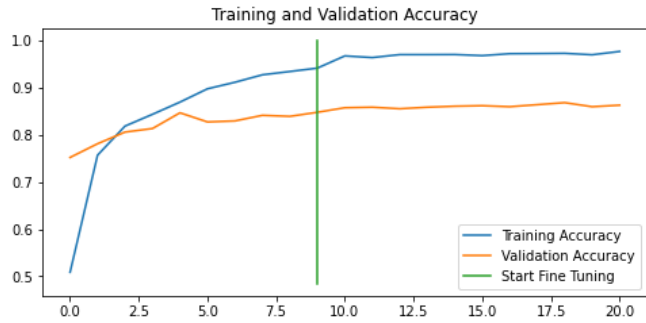
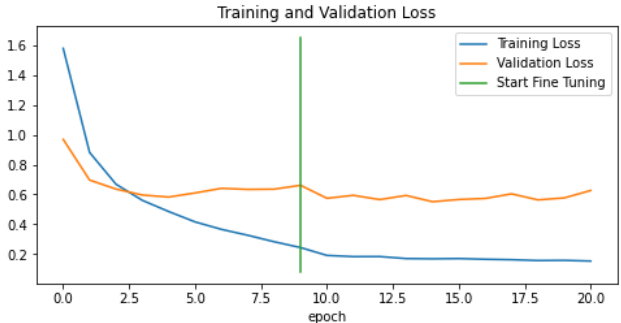
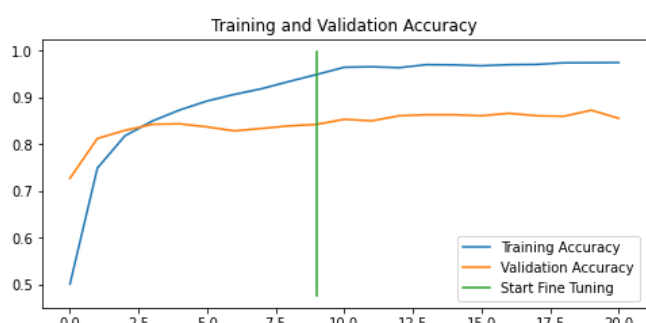
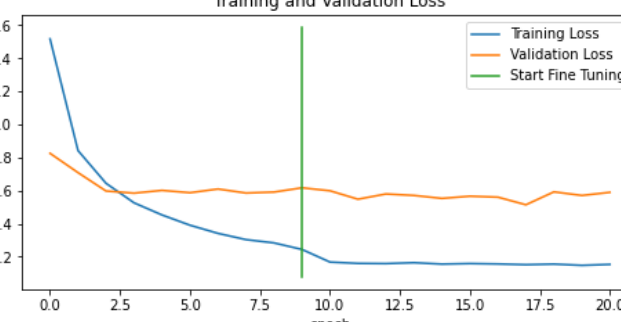
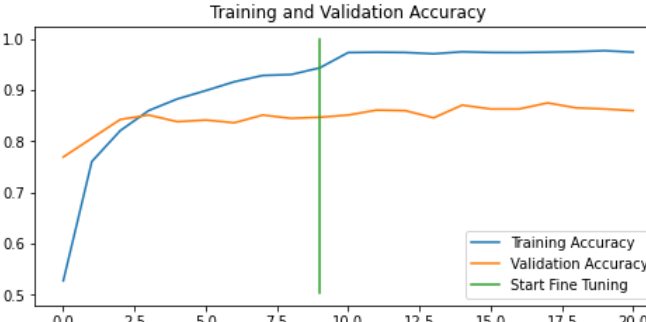
در جدول شماره ۲ تعداد نورون‌ها با ۳ لایه که از بخش قبل بهترین دقت را داشت تغییر داده شده است. ابتدا تعداد نورون‌های هر لایه را نصف کرده‌ایم و یک بار نیز تعداد نورون‌های هر لایه را ۲ برابر کرده‌ایم. در هر دو حالت دقت نسبت به حالت اولیه کاهش یافته است بنابراین همان ۱۰۶۴ و ۱۲۸ نورون بهترین دقت را دارد.

بخش دوم: بررسی تعداد لایه‌هایی که وزن آن‌ها را ثابت در نظر می‌گیریم

در جدول شماره ۳ تعداد لایه‌هایی که وزن آن‌ها را ثابت می‌گیریم با ۳ مقدار مختلف ۳۰۰ و ۳۵۰ و ۴۰۰ بررسی شده است. همانطور که مشاهده می‌شود تعداد ۳۰۰ بیشترین دقت را دارد و بعد از آن به ترتیب ۳۵۰ و ۴۰۰ لایه. بنابراین با افزایش تعداد لایه‌هایی که tune می‌کنیم دقت افزایش پیدا می‌کند (البته با در نظر گرفتن این نکته که لایه‌های خیلی ابتدایی را tune نمی‌کنیم چون فیچرهای کلی را استخراج می‌کنند).

بخش سوم: مقایسه

در تمرین قبل بهترین دقت به دست آمده بر روی داده‌های تست برابر ۵۵ درصد بود در حالی که در این تمرین با استفاده از شبکه پیش آموزش دیده در بهترین حالت دقت روی داده‌های تست به ۹۰ درصد رسید. همچنین آموزش شبکه از پیش آموزش دیده با تعداد epoch های بیشتری انجام شد زیرا زمان کمتری نیاز داشت. بنابراین شبکه از پیش آموزش دیده هم از نظر سرعت و هم از نظر دقت و عملکرد برای این مسئله بهتر است.

نمودار خطا	نمودار دقت	خطا	دقت	تعداد لایه‌های با وزن ثابت
		0.3970	0.9020	300
		0.4192	0.8972	350
		0.4508	0.8884	400

جدول ۳ بررسی تعداد لایه‌های با وزن ثابت