تمرین دوم پایتون
شبکه‌های عصبی بهار ۱۴۰۲

آریا ابراهیمی ۹۸۲۲۷۶۲۱۷۵

هدف از این تمرین، بررسی انتقال دانش در شبکه‌های عصبی عمیق می‌باشد. برای این کار از سه شبکه VGG16، ResNet50 و DenseNet121 آموزش داده شده بر روی ImageNet استفاده می‌کنیم. شبکه های یادگیری شده در لایه ابتدایی شامل فیچر های ابتدایی تر می‌شوند و هرچقدر عمیق‌تر می‌شوند، ویژگی های پیچیده‌تری را شامل می‌شوند. هدف این است تا با استفاده از وزن های یادگیری شده لایه های ابتدایی با تعداد داده کمتر و سرعت بیشتری بتوانیم classification را انجام دهیم.

در این تمرین با استفاده از داده های سگ و گربه، عمل‌کرد شبکه های معرفی شده را با درصد های متفاوت لایه های freeze شده بررسی می‌کنیم. در نهایت ویژگی های یادگیری شده مربوط به داده‌ها را به یک RandomForestClassifier میدهیم و تسک نهایی مربوط به classification را با استفاده از آن نیز انجام می‌دهیم و با classification انجام شده توسط تک نورون لایه آخر شبکه مقایسه می‌کنیم.

۱- بخش اول

همان‌طور که در جدول۱ مشاهده می‌شود، هرچقدر تعداد لایه های trainable بیشتر شود، میزان accuracy مدل کمتر می‌شود و بهترین عمل‌کرد مربوط به تعداد لایه های قابل یادگیری ۱۰ درصد و ۳۰ درصد می‌باشد و این برای هر سه شبکه VGG16 و ResNet50 و DenseNet121 برقرار است. می‌توان نتیجه گرفت که انتقال دانش می‌تواند به یادگیری بهتر کمک کند و اگر داده زیادی برای آموزش یک شبکه نداریم، استفاده از یک شبکه از پیش آموزش داده شده می‌تواند از overfit شدن و مشکلات دیگر جلوگیری کند.

جدول ۱: انتقال دانش از مدل های گفته شده بر روی یک batch از دیتاست سگ و گربه.

Model	VGG16					ResNet50					DenseNet121				
Trainable	10%	30%	50%	70%	100%	10%	30%	50%	70%	100%	10%	30%	50%	70%	100%
Train Accuracy	94.40	94.50	92.59	75.14	82.09	96.95	96.69	96.45	96.30	96.15	97.66	97.42	96.98	97.21	96.73
Validation Accuracy	97.78	97.75	96.84	84.45	86.38	98.91	98.58	97.92	98.64	98.09	98.98	99.08	99.05	99.07	98.15

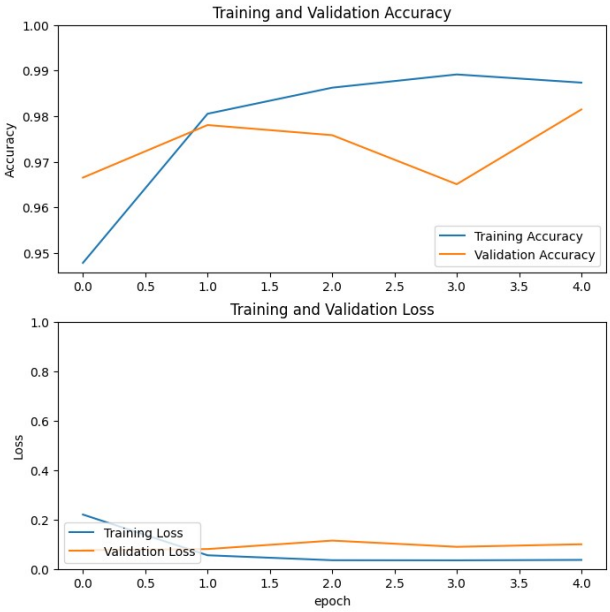
همان‌طور که مشاهده می‌شود در مجموع DenseNet121 عملکرد بهتری نسبت به دو مدل دیگر دارد و در هر مدل هم لایه های قابل یادگیری بین ۱۰درصد تا ۳۰درصد بهترین نتیجه را نمایش میدهند.

سوالی که پیش می‌آید این است که آیا اگر تمامی لایه های مدل اصلی را freeze کنیم نسبت به نتایج فعلی خروجی بهتری خواهیم داشت؟ نتایج به دست آمده برای زمانی که تمامی لایه ها را freeze کنیم در جدول ۲ قابل مشاهده است.

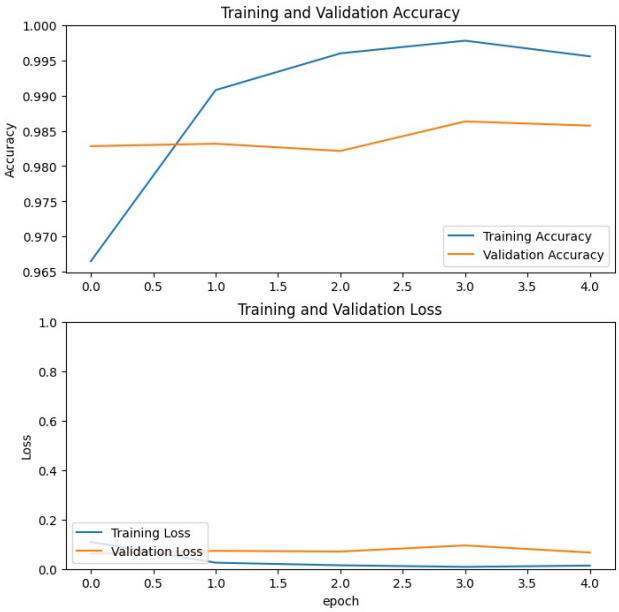
جدول ۲: انتقال دانش زمانی که تمامی لایه ها freeze شوند

Model	VGG16	ResNet50	DenseNet121
Train Accuracy	94.7	96.95	97.54
Validation Accuracy	97.53	98	98.36

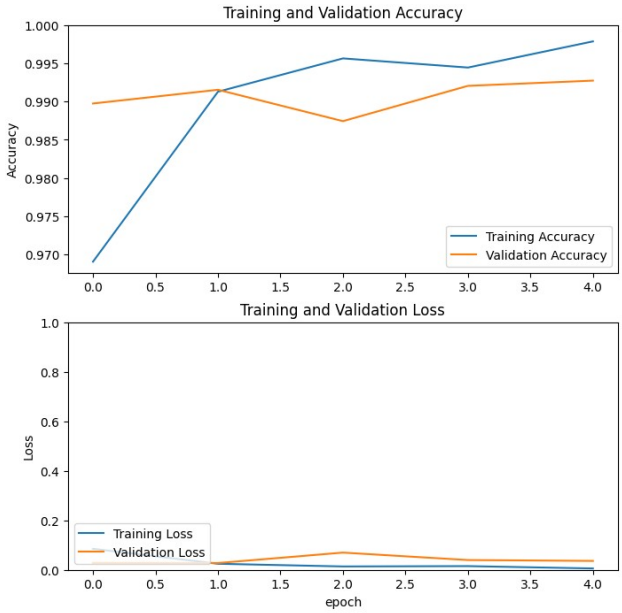
همان‌طور که مشاهده می‌شود در حالتی که تمامی لایه های شبکه های گفته شده freeze شوند، نتیجه بهتری مشاهده نمی‌شود.



شکل ۱: انتقال دانش با استفاده از مدل VGG16 و ۳۰ درصد لایه های trainable در پنج epoch



شکل ۲: انتقال دانش با استفاده از مدل ResNet50 و ۳۰ درصد لایه های trainable در پنج epoch



شکل ۳: انتقال دانش با استفاده از مدل DenseNet121 و ۳۰ درصد لایه های trainable در پنج epoch

۲- بخش دوم

در این بخش، ویژگی های یادگیری شده مربوط به نسخه ۳۰ درصد شبکه ها را برای همان دیتاست استخراج کرده و با استفاده RandomForestClassifier کتابخانه SciKit آن‌ها را دسته‌بندی می‌کنیم. چالشی که در قسمت پیاده سازی این بخش وجود دارد این است که دیتاست را به صورت BatchDataset دریافت کرده ایم (شامل داده های یادگیری و لیبل مربوط به آنها) که باعث می‌شود به راحتی به داده و لیبل دسترسی نداشته باشیم و از طریق متد take میتوانیم روی batch ها پیمایش کنیم. یک راه میتواند این باشد که همزمان با پیمایش روی داده‌ها، آن‌ها را به همراه لیبل متناظرشان ذخیره کنیم. این روش رم زیادی اشغال می‌کند و باعث کرش کردن محیط Google Colab می‌شود.

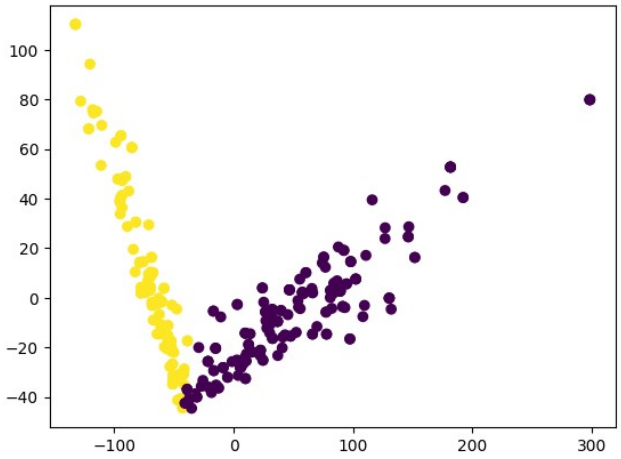
راه دیگر این است که RandomForestClassifier را همزمان با پیمایش روی دیتاست به صورت mini-batch آموزش دهیم. با این کار دیگر نیازی به ذخیره داده ها نیست ولی زمان بیشتری نیاز دارد.

دقت RandomForestClassifier برای داده های validation برای سه شبکه در نسخه ۳۰ درصد در جدول ۳ قابل مشاهده است.

جدول ۳: دقت بر روی دیتاست validation با استفاده از RandomForestClassifier

Model	VGG16	ResNet50	DenseNet121
Validation Accuracy	97.45	98.56	99.25

همان‌طور که مشاهده می‌شود، طبقه بندی با استفاده از RandomForestClassifier روی ویژگی های استخراج شده از شبکه ها نتایج برابر و حتی در مواردی بهتر از classification خود شبکه ها ارائه می‌دهد.



شکل ۴: ویژگی های کاهش dimensionality داده شده توسط PCA و یادگیری شده توسط شبکه (قبل از اعمال تابع فعال‌ساز relu)

می‌توان نتیجه گرفت که مهم‌ترین کاری که شبکه ها انجام میدهند این است که ویژگی یاد می‌گیرند و شاید تسک classification را طبقه‌بندی مثل RandomForestClassifier بهتر انجام دهد ولی چیزی که شبکه های عصبی را متمایز می‌کند و باعث می‌شود خوب عمل کند، ویژگی های یادگیری شده خوب است.