



دانشکده مهندسی
گروه مهندسی کامپیوتر
مقطع کارشناسی

عنوان

CNN Fine Tuning

نگارش

کیارش آستابوس

اردیبهشت ۱۴۰۴

چکیده

در این پروژه ابتدا دو معماری ResNet-۱۸ و DenseNet-۱۲۱ بر روی مجموعه داده Flowers-۱۰۲ با چهار سطح مختلف از درصد پارامترهای قابل تعلیم (۸۰٪، ۵۰٪، ۳۰٪ و تنها لایه طبقه‌بندی) فاین‌تیون شدند و برای هر سناریو بهترین هایپرپارامترها با جستجو تعیین گردید. سپس دقت نهایی هر ترکیب روی مجموعه ولیدیشن و تست اندازه‌گیری و معماری برتر (DenseNet-۱۲۱ با ۸۰٪ پارامترهای قابل تعلیم) برای فاز دوم انتخاب شد. در فاز دوم، برای مطالعه تأثیر اجزای فرکانسی تصویر، تبدیل فوریه دوبعدی روی تصاویر اعمال و فیلترهای low-pass و high-pass با شعاع‌های مختلف $r \in \{10, 30, 50, 70, 90\}$ استفاده شد؛ دقت مدل منتخب بر روی تصاویر فیلترشده محاسبه و نشان داده شد که حذف بیش از حد فرکانس‌های بالا یا پایین به ترتیب منجر به افت عملکرد می‌شود.

فهرست مطالب

۱	۱ فاز اول: پیدا کردن مدل منتخب
۱	۱-۱ رویکرد
۱	۲-۱ نتایج
۲	۳-۱ مقایسه معماری‌ها
۲	۴-۱ مدل منتخب
۲	۱-۴-۱ هایپرپارامترها
۳	۲-۴-۱ نمودارهای Loss و Accuracy
۴	۲ فاز دوم: اثر فیلترهای فرکانسی بر دقت مدل منتخب
۴	۱-۲ رویکرد
۴	۲-۲ تأثیر بصری فیلترهای Low-Pass و High-Pass بر روی تصویر
۶	۳-۲ نتایج
۶	۱-۳-۲ دقت و زیان
۷	۲-۳-۲ دقت بر حسب r
۹	۳ نتیجه‌گیری

فهرست جداول

- ۱-۱ دقت نهایی مدل‌ها بر حسب درصد پارامترهای قابل تعلیم ۲
- ۲-۱ مقادیر نهایی هایپرپارامترهای مدل DenseNet-۱۲۱ (۸۰٪ پارامتر تعلیم شده) . . . ۳

فهرست تصاویر

۳	۱-۱ مقایسه منحنی‌های Accuracy و loss برای داده‌های Train و Test
۵	۱-۲ تأثیر بصری فیلتر
۷	۲-۲ نمودارهای Loss و Accuracy
۷	۳-۲ دقت مدل منتخب بر حسب پارامتر r در فیلترهای فرکانسی

فصل ۱

فاز اول: پیدا کردن مدل منتخب

۱-۱ رویکرد

برای هر یک از دو ساختار ۱۸-ResNet و ۱۲۱-DenseNet چهار سناریوی فریز/آزادسازی پارامترها پیاده‌سازی شد:

۱. ۸۰٪ پارامترها قابل تعلیم، ۲۰٪ فریز

۲. ۵۰٪ پارامترها قابل تعلیم، ۵۰٪ فریز

۳. ۳۰٪ پارامترها قابل تعلیم، ۷۰٪ فریز

۴. فقط لایه طبقه‌بندی نهایی قابل تعلیم، سایر پارامترها فریز

برای هر سناریو، جستجوی بهترین هایپرپارامتر (شامل نرخ یادگیری هر دسته از لایه‌ها و تعداد لایه‌های طبقه‌بند و Optimizer) انجام شد و بهترین مدل بر اساس بیشینه دقت بر روی مجموعه اعتبارسنجی انتخاب گردید. برای دسته‌بندی لایه‌های قابل تعلیم، در تمامی حالت‌ها به ۲ گروه تقسیم‌بندی شده‌اند.

۲-۱ نتایج

دقت نهایی (بر روی مجموعه ولیدیشن) هر سناریو در جدول ۱-۱ آورده شده است.

با افزایش تعداد پارامترهای قابل تعلیم، ظرفیت مدل برای تطبیق با ویژگی‌های خاص مجموعه ۱۰۲ Flowers-افزایش می‌یابد و از این‌رو دقت بهبود می‌یابد. در سناریوی تنها طبقه‌بند، بخش عمده دانش

استخراج شده از پیش‌آموزش حفظ می‌شود اما توانایی تطبیق با بافت‌های جدید کاهش می‌یابد که منجر به پایین‌ترین دقت می‌شود. کاهش تدریجی دقت از ۸۰٪ به ۳۰٪ تعلیم نشان‌دهنده کاهش ظرفیت مدل برای بازنمایی پیچیدگی‌های داده است.

جدول ۱-۱: دقت نهایی مدل‌ها بر حسب درصد پارامترهای قابل تعلیم

مدل	۸۰٪ تعلیم	۵۰٪ تعلیم	۳۰٪ تعلیم	تنها طبقه‌بند
ResNet-۱۸	۹۱.۶٪	۸۹.۵٪	۸۷.۴٪	۸۴.۹٪
DenseNet-۱۲۱	۹۲.۴٪	۹۱.۴٪	۹۱.۸٪	۸۹.۶٪

برای صرفه‌جویی در فضای گزارش، فقط نمودارهای Accuracy و Loss برای مدل منتخب بررسی شده. به ازای هر پارامتر، (۱۲۸ حالت) نمودار دقت و زیان در نوت‌بوک وجود دارد.

۳-۱ مقایسه معماری‌ها

در جدول ۱-۱ دقت دو مدل در سناریوی ۸۰٪ تعلیم نمایش داده شده است.

شبکه DenseNet-۱۲۱ با اتصال‌های چگال (dense connections) جریان گرادین را تسهیل می‌نماید و امکان استفاده بهتر از ویژگی‌های استخراج شده در لایه‌های عمیق را فراهم می‌کند. این امر منجر به بازنمایی غنی‌تر و در نتیجه دقت بالاتر نسبت به ResNet-۱۸ شد. در حالی که ResNet-۱۸ با residual blocks عملکرد مناسبی ارائه داد، ظرفیت کمتر آن در استخراج ویژگی‌های ظریف گل‌ها باعث کاهش دقت گردید.

۴-۱ مدل منتخب

۱-۴-۱ هایپرپارامترها

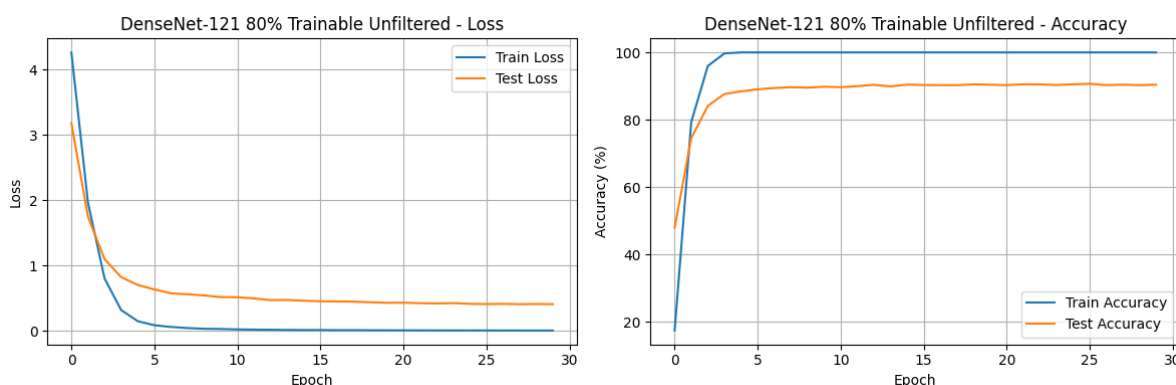
در نهایت پارامترهای مدل منتخب طبق جدول ۲-۱ است.

جدول ۱-۲: مقادیر نهایی هایپرپارامترهای مدل DenseNet-۱۲۱ (۸۰٪ پارامتر تعلیم شده)

مقدار	هایپرپارامتر
Adam	بهینه‌ساز (Optimizer)
1×10^{-4}	نرخ یادگیری لایه‌های پایه (base layers)
1×10^{-3}	نرخ یادگیری لایه طبقه‌بند نهایی
۳۲	Batch size

۱-۴-۲ نمودارهای Loss و Accuracy

نمودارهای دقت و زیان در شکل ۱-۱ قابل مشاهده هستند.



شکل ۱-۱: مقایسه منحنی‌های Accuracy و loss برای داده‌های Train و Test

منحنی‌های Loss و Accuracy در شکل ۱-۱ نشان می‌دهند که خطا روی مجموعه آموزش به سرعت در epoch ۵ اول کاهش یافته و سپس حول یک مقدار کم تثبیت شده است. خطای مجموعه تست نیز کاهش مشابهی را دنبال می‌کند و پس از epoch حدوداً ششم به شیب بسیار کمی کاهش می‌یابد. نمودار دقت هم همین رفتار را دارد و در نهایت دقت داده‌های تست برابر ۹۰.۴٪ و دقت داده‌های آموزش برابر ۱۰۰٪ است.

فصل ۲

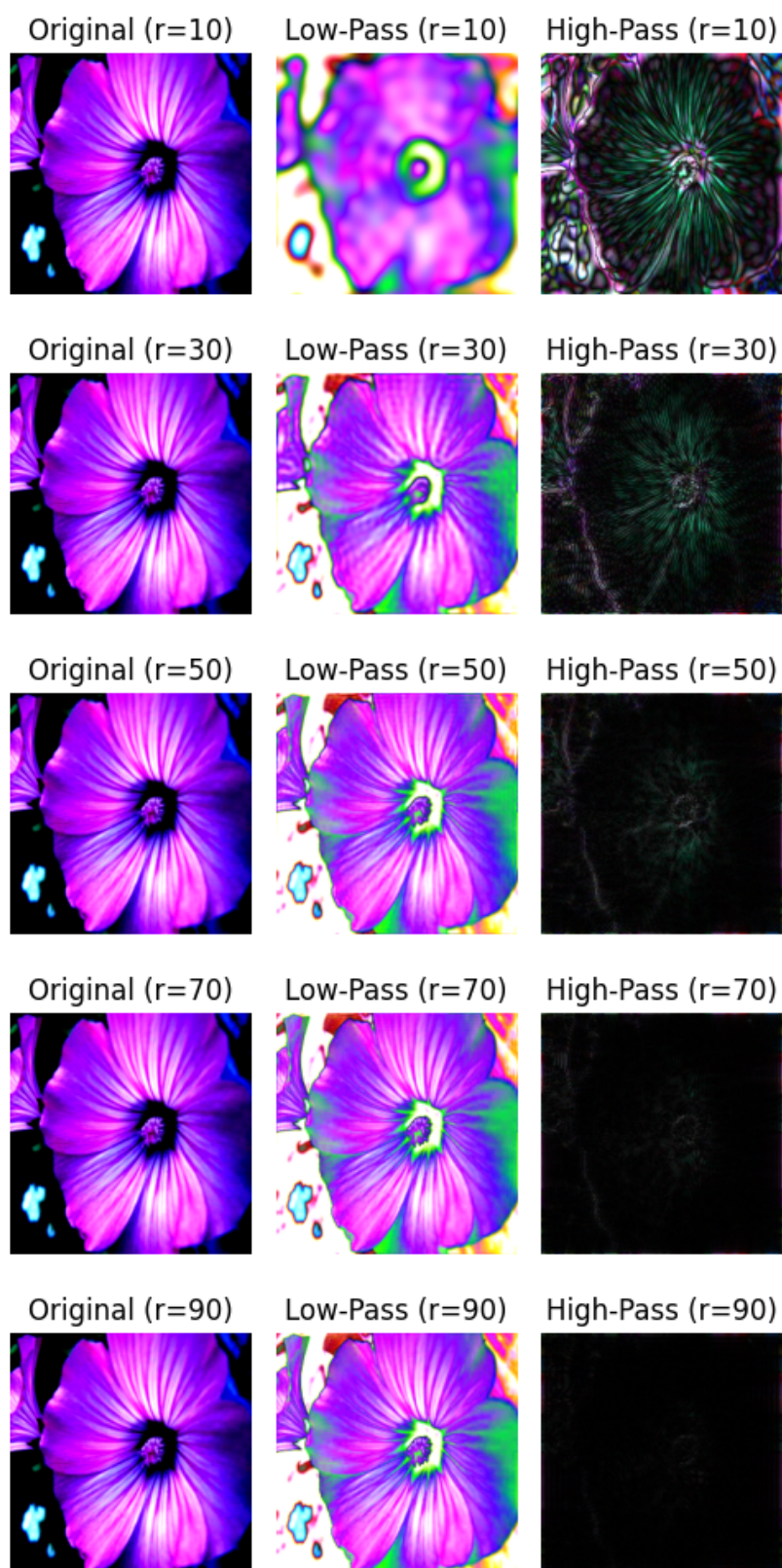
فاز دوم: اثر فیلترهای فرکانسی بر دقت مدل منتخب

۱-۲ رویکرد

مدل منتخب از فاز اول (DenseNet-۱۲۱ با ۸۰٪ پارامترهای قابل تعلیم) برای ارزیابی روی تصاویر فیلترشده با روش FFT استفاده شد. فیلترهای low-pass و high-pass با شعاعهای مختلف $r \in \{10, 30, 50, 70, 90\}$ در حوزه فرکانس اعمال گردیدند. برای هر مقدار، تصاویر بازسازی و دقت مدل روی آنها اندازه گیری شد.

۲-۲ تأثیر بصری فیلترهای Low-Pass و High-Pass بر روی تصویر

برای درک بهتر تأثیر فیلترهای فرکانسی بر داده ورودی، تصاویر نمونه ای از یک گل با اعمال فیلترهای Low-Pass و High-Pass با شعاعهای مختلف $r \in \{10, 30, 50, 90\}$ تهیه شده اند. نتایج تصویری این فیلترها در شکل ۱-۲ نمایش داده شده است.



شکل ۲-۱: تأثیر بصری فیلتر

با کاهش شعاع r ، در فیلتر Low-Pass مؤلفه‌های فرکانسی بالا (لبه‌ها، جزئیات بافت و خطوط دقیق) به تدریج حذف می‌شوند. همان‌طور که در شکل ۱-۲ دیده می‌شود، برای $r = 90$ تصویر تفاوت زیادی با نسخه اصلی ندارد؛ اما در $r = 30$ جزئیات گل محو شده و در $r = 10$ تنها ساختار کلی رنگی و شکل کلی گل قابل تشخیص است. کاهش وضوح و حذف جزئیات باعث می‌شود مدل دقت کمتری در تشخیص دقیق کلاس گل‌ها داشته باشد.

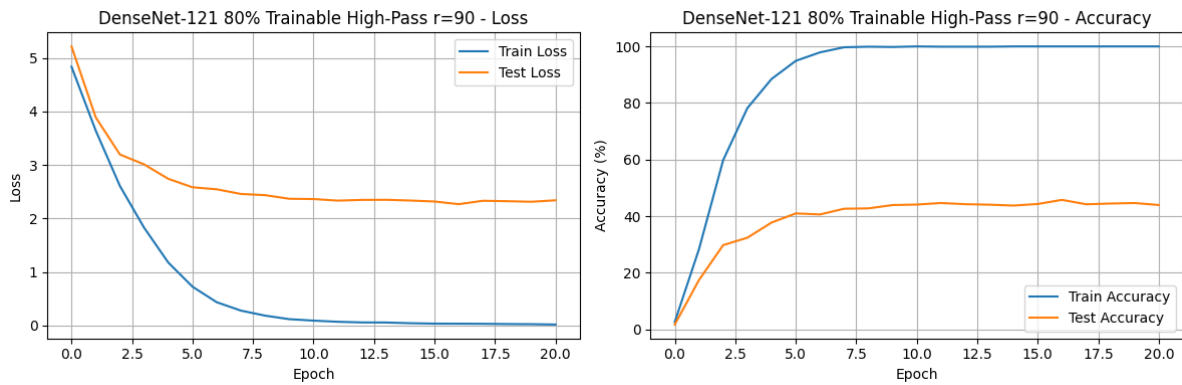
در فیلتر High-Pass، با کاهش r مؤلفه‌های فرکانس پایین (ساختار کلی، رنگ غالب و شکل کلی) حذف شده و تنها جزئیات ریز، لبه‌ها و نویز حفظ می‌شود. در شکل ۱-۲ برای $r = 90$ تصویر هنوز تا حدودی قابل درک است؛ اما در $r = 30$ تنها خطوط مشخص و لبه‌های گل باقی مانده‌اند و در $r = 10$ تصویر تقریباً غیرقابل تشخیص است. حذف ساختار کلی در این حالت، مدل را از اطلاعات پایه مورد نیاز برای تشخیص بازمی‌دارد و باعث افت شدید عملکرد می‌شود.

از نظر بصری، فیلتر Low-Pass در مقادیر r پایین‌تر موجب محو شدن اطلاعات مهم و کاهش قابلیت تشخیص می‌شود؛ در حالی که فیلتر High-Pass باعث حذف بافت‌های عمومی شده و تصویر را برای انسان و مدل غیرقابل تفسیر می‌کند.

۳-۲ نتایج

۱-۳-۲ دقت و زیان

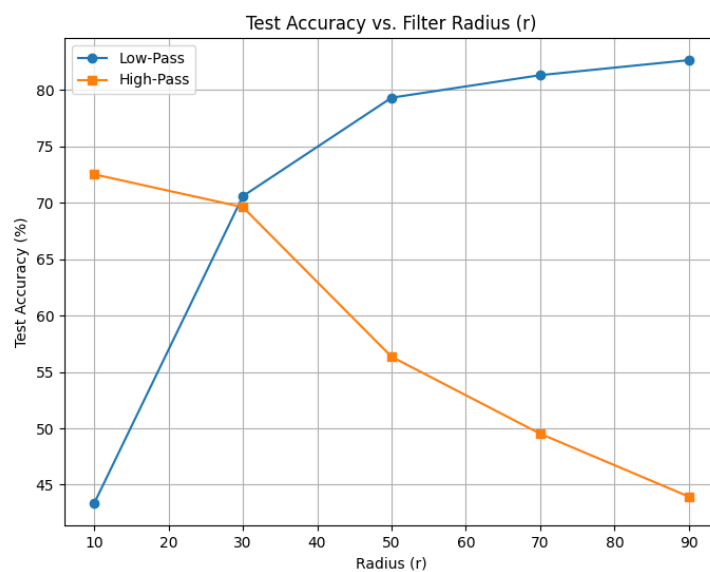
یک نمونه از نمودار Accuracy و Loss هنگام آموزش مدل در شکل ۲-۲ قابل مشاهده است. رفتار مدل در تمامی نمودارها به همین صورت است که هر ۲۰ نمودار در فایل نوت‌بوک قابل مشاهده هستند. تنها تفاوت آن‌ها در دقت داده‌های تست است و همگی دقت ۱۰۰٪ بر روی داده‌های آموزش داشته‌اند. با توجه به اختلاف زیاد دقت داده‌های تست و آموزش، نتیجه گرفته می‌شود که مدل‌ها overfit شده‌اند.



شکل ۲-۲: نمودارهای Loss و Accuracy

۲-۳-۲ دقت بر حسب r

در شکل ۲-۳ نمودار دقت مدل بر حسب مقدار r برای دو نوع فیلتر نمایش داده شده است.



شکل ۲-۳: دقت مدل منتخب بر حسب پارامتر r در فیلترهای فرکانسی

اثر فیلتر **Low-Pass** با کاهش مقدار r ، اجزای فرکانس بالا (جزئیات ریز تصویر) حذف شده و مدل تنها روی ساختار کلی تصویر (بافت عمومی گل) عمل می‌کند. دقت تا $r = 50$ تقریباً ثابت ماند اما برای $r < 30$ کاهش چشمگیری مشاهده شد؛ زیرا برخی ویژگی‌های ضروری برای تمایز گونه‌ها از بین رفته‌اند.

High-Pass اثر فیلتر **High-Pass** فیلترهای High-Pass جزئیات ریز (لبه‌ها و بافت‌های دقیق) را نگه می‌دارند ولی ساختار کلی را حذف می‌کنند. در این حالت دقت به طور پیوسته با کاهش r (نگه‌داشتن اجزای فرکانس بالاتر) کاهش یافت و نشان می‌دهد که ویژگی‌های فرکانس پایین برای تشخیص گل‌ها اهمیت بیشتری دارند.

فصل ۳

نتیجه‌گیری

طی نتایج به دست آمده نتیجه گرفته می‌شود که:

- افزایش درصد پارامترهای قابل تعلیم منجر به بهبود دقت می‌شود؛ زیرا انعطاف بیشتری برای یادگیری ویژگی‌های جدید فراهم می‌آورد.
- معماری DenseNet-۱۲۱ به دلیل اتصالات چگال، عملکرد بهتری نسبت به ResNet-۱۸ ارائه کرد.
- فیلترهای Low-Pass تا حد معینی (حدود $r = ۵۰$) تأثیر کمی بر دقت دارند، اما حذف بیش از حد جزئیات ریز موجب افت عملکرد می‌شود. فیلترهای High-Pass به تنهایی برای این مسئله مناسب نیستند.