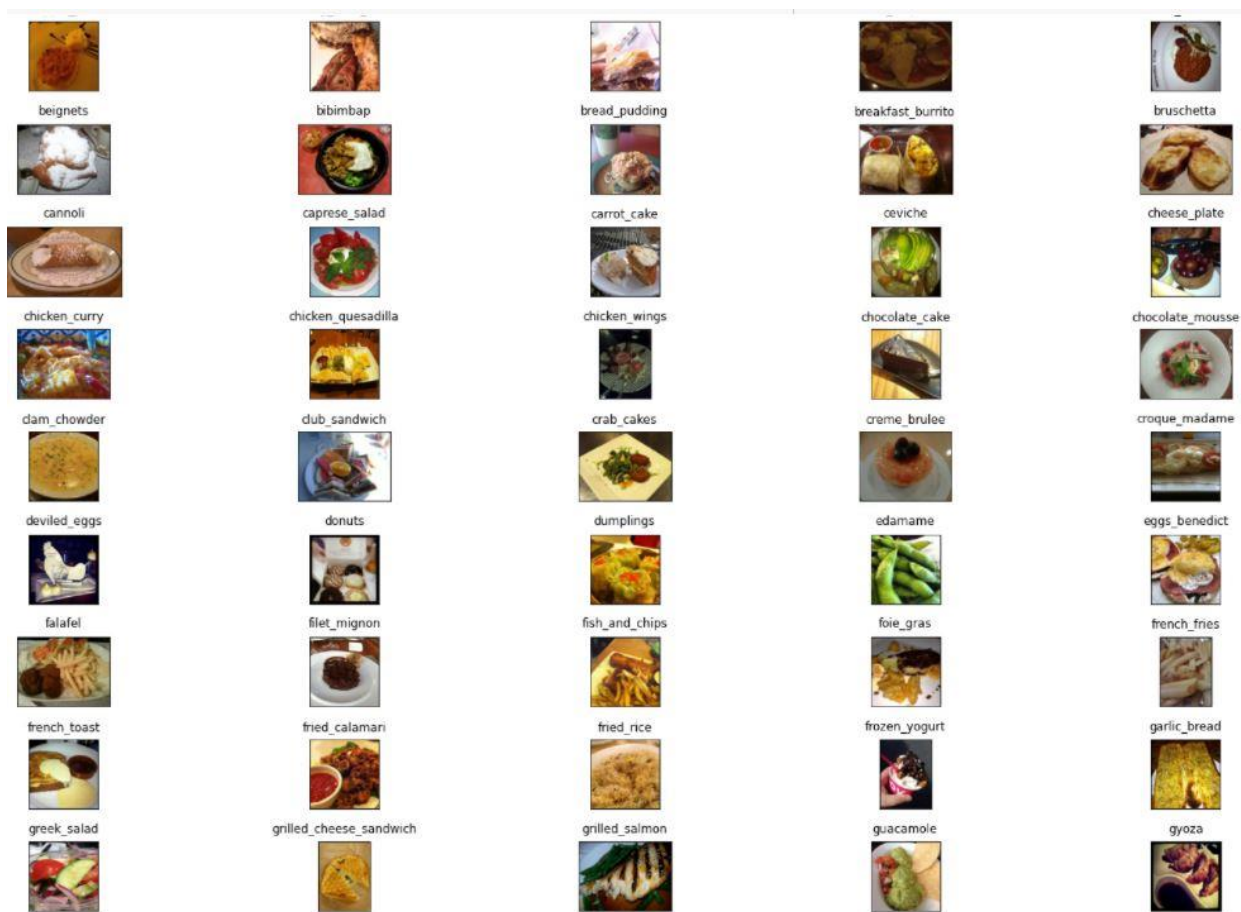


تمرین چهارم درس شبکه های عصبی

آماده سازی داده ها

در این تمرین قصد داریم با استفاده از شبکه های کانولوشنی، تصاویر غذاهای مختلف را دسته بندی نماییم. ابتدا مجموعه داده های food101 رو از tensorflow بارگذاری می کنیم. این مجموعه شامل ۱۰۱ کلاس یا ۱۰۱ نوع غذای مختلف است که برای هر کلاس ۱۰۰۰ داده وجود دارد که ۷۵۰ تای آن داده های آموزش و ۲۵۰ تای آن داده های تست می باشند. تصاویر آموزشی شفاف نیستند و دارای نویز می باشند. این کار به طور عامدانه صورت گرفته تا از overfitting مدل جلوگیری شود. اما داده های تست بدون نویز و clean هستند. علاوه بر داده ها یک پوشه meta نیز در کنار داده ها وجود دارد. این پوشه شامل ۳ فایل train.txt، test.txt و class.txt می باشد که به ترتیب مشخص می کند کدام تصاویر به مجموعه آموزش، مجموعه تست و کدام کلاس تعلق دارد. در اینجا تعدادی از تصاویر را به صورت رندوم انتخاب می کنیم و با برچسب آن به نمایش می گذاریم.



شکل ۱: برخی از تصاویر مجموعه داده ها به همراه برچسب آن ها

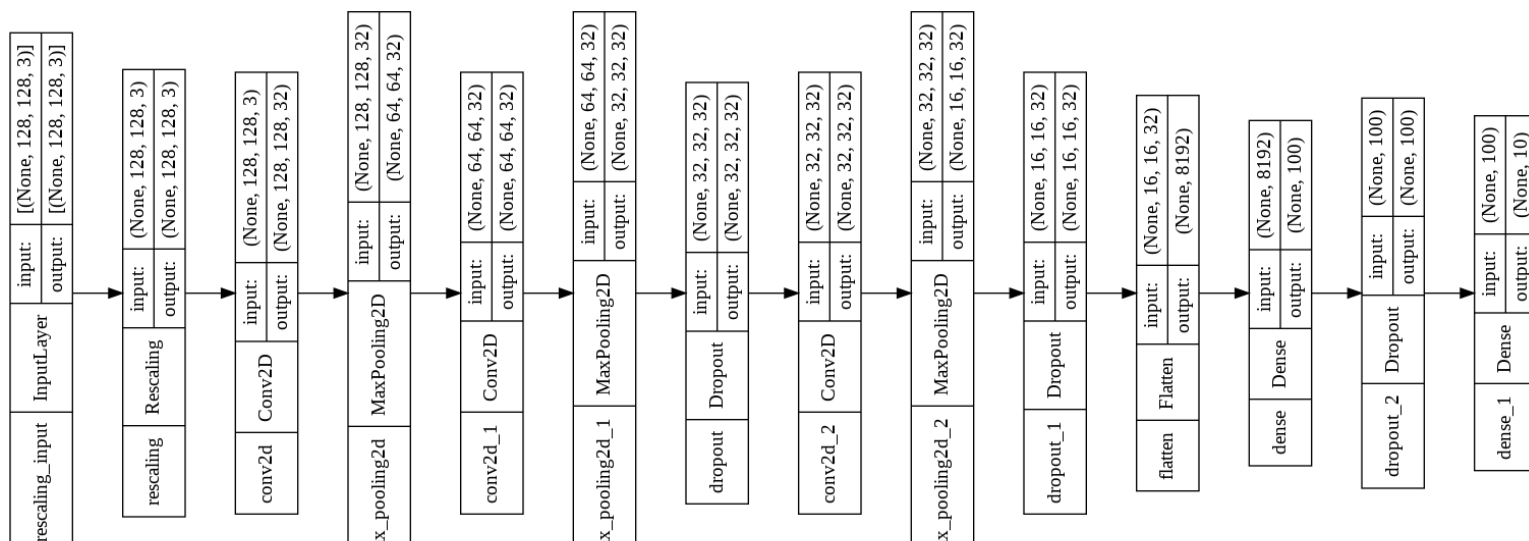
در ادامه به کمک محتویات پوشه meta داده‌ها را به مجموعه‌های آموزش و تست برای هر کلاس تقسیم می‌کنیم. ما اکنون داده‌های آموزش و تست را آماده کرده‌ایم، اما برای آزمایش معماری‌های مختلف، کار روی کل داده‌ها با ۱۰۱ کلاس زمان و محاسبات زیادی می‌برد. برای انجام آزمایش مورد نظر، مجموعه داده را به ۱۰ کلاس محدود می‌کنم و train_min و test_mini را ایجاد می‌کنیم. توجه شود که برای انتخاب ۱۰ کلاس از ۱۰۱ کلاس جایگشت‌های مختلفی وجود دارد که هر یک از این جایگشت‌ها بسته به این که ۱۰ کلاس انتخاب شده با هم چقدر همبستگی دارند روی مقدار accuracy تاثیر می‌گذارد. این مقدار می‌تواند حدود ۱۰ درصد باشد. در اینجا ما ۱۰ کلاس را به صورت رندوم انتخاب می‌کنیم تا مدل را نسبت به کلاس‌های خاص بایاس نکنیم.

طراحی شبکه کانولوشنی

در این قسمت می‌خواهیم به ارزیابی ساختارهای مختلف یک شبکه کانولوشنی که بعد از هر لایه کانولوشنی یک لایه زیرنمونه برداری وجود دارد بپردازیم. برای این کار ابتدا به بررسی تعداد جفت‌های مختلف شامل یک لایه کانولوشنی و یک لایه زیرنمونه برداری می‌پردازیم. به کمک ImageDataGenerator ابعاد تصاویر را 128×128 می‌کنیم. همچنین برای آموزش شبکه از ۳۰ تکرار با early stop=4 استفاده شد که در صورتی که مدل پس از ۴ تکرار بهبودی نداشته باشد روند آموزش متوقف می‌شود.

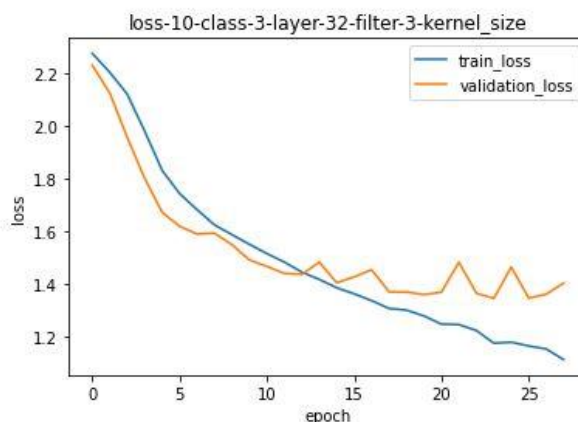
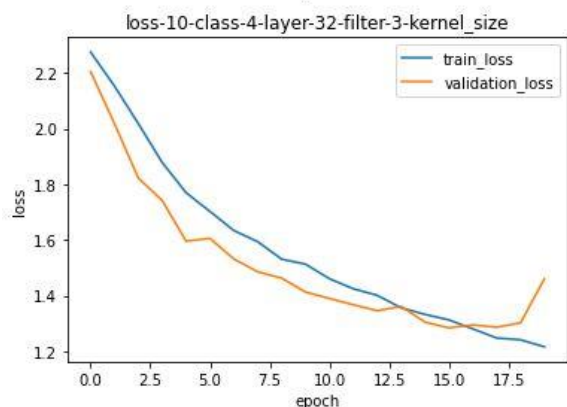
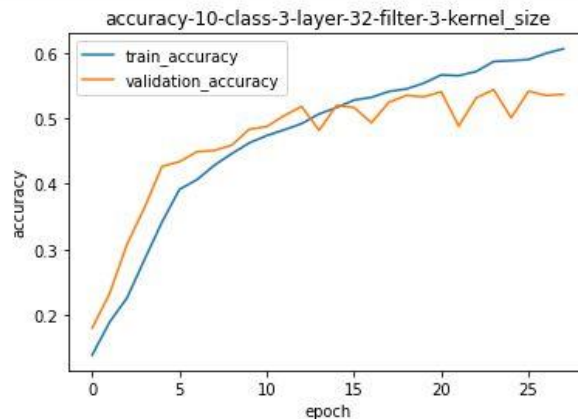
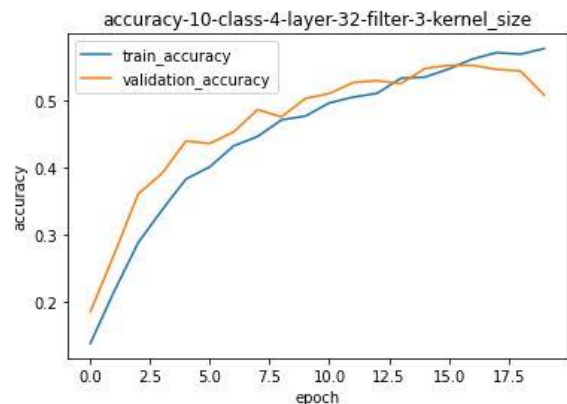
تعداد جفت‌های مختلف

ساختار شبکه به این صورت است. ابتدا یک لایه rescaling استفاده می‌کنیم که مقدار همه پیکسل‌ها را بین ۰ تا ۱ بیاوریم. سپس به تعداد مختلف در هر آزمایش جفت‌های کانولوشن-زیر نمونه برداری را قرار می‌دهیم هر یک لایه‌های کانولوشن شامل ۳۲ کرنل به سایز 3×3 می‌باشد. همچنین، برای زیر نمونه برداری از MaxPooling دو بعدی با ابعاد 2×2 استفاده می‌کنیم. به جز جفت اول بعد از هر جفت یک dropout به مقدار ۱۰ درصد قرار می‌دهیم تا به generalization مدل کمک کند. در همه لایه‌های کانولوشن از تابع فعال‌سازی relu استفاده می‌کنیم. همچنین با قرار دادن padding به مقدار same به مقداری که لازم است در دو بعد به تصویر ورودی صفر اضافه می‌کنیم تا فیلتر به همه ورودی‌ها اعمال شود. پس از بخش کانولوشن، برای بخش دسته‌بندی اط دو لایه دنس به ترتیب ۱۰۰ و ۱۰ نرون و توابع فعال‌سازی relu و softmax استفاده می‌کنیم. در زیر ساختار شبکه را برای ۳ جفت مشاهده می‌کنید.



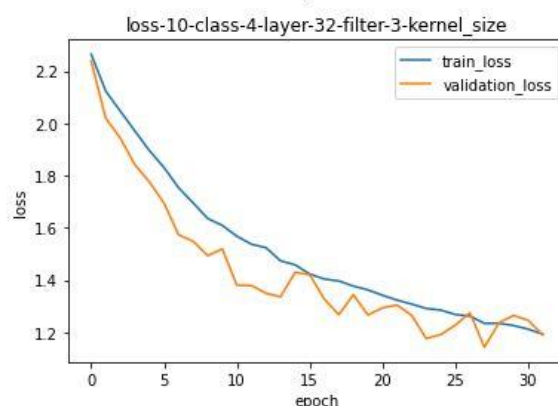
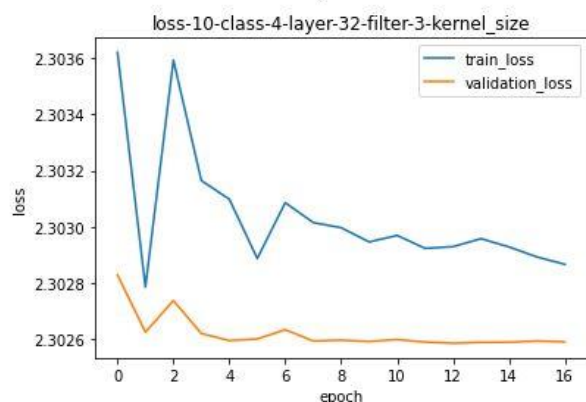
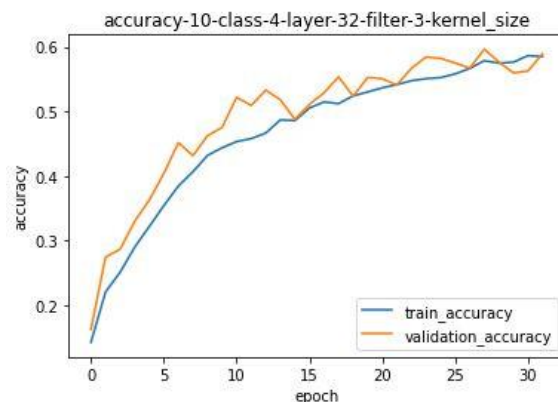
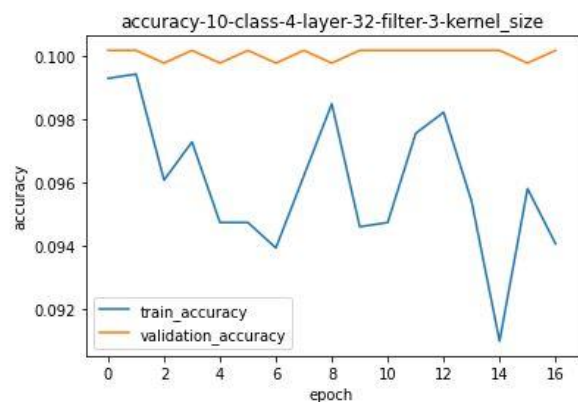
شکل ۱: ساختار شبکه برای ۳ جفت لایه کانولوشنی

این تست را برای سه جفت تا ۷ جفت انجام دادیم. باتوجه به اینکه ابعاد تصاویر ورودی را 128×128 کردیم و هر لایه pooling 2×2 ابعاد خروجی را در هر بعد نصف می‌کند بنابراین استفاده از بیش از ۷ جفت امکان پذیر نیست. در زیر نمودارهای مربوط به accuracy و loss برای هر حالت به نمایش گذاشته شده است.



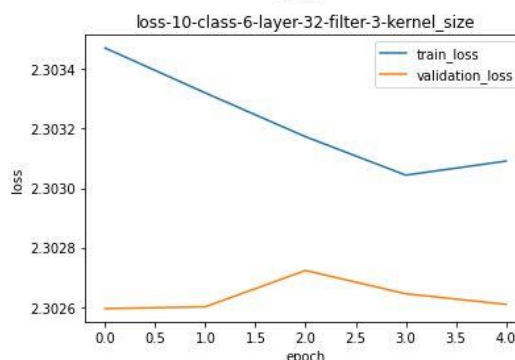
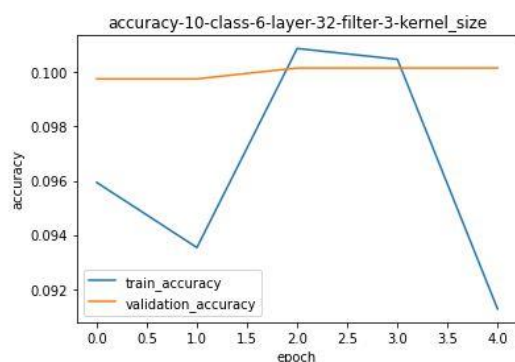
شکل ۳: نمودارهای accuracy و loss برای شبکه با ۴ جفت هر یک با ۳۲ کرنل به سایز 3×3

شکل ۲: نمودارهای accuracy و loss برای شبکه با ۳ جفت هر یک با ۳۲ کرنل به سایز 3×3



شکل ۵: نمودارهای **accuracy** و **loss** برای شبکه با ۶ جفت هر یک با ۳۲ کرنل به سائز ۳*۳

شکل ۴: نمودارهای **accuracy** و **loss** برای شبکه با ۵ جفت هر یک با ۳۲ کرنل به سائز ۳*۳

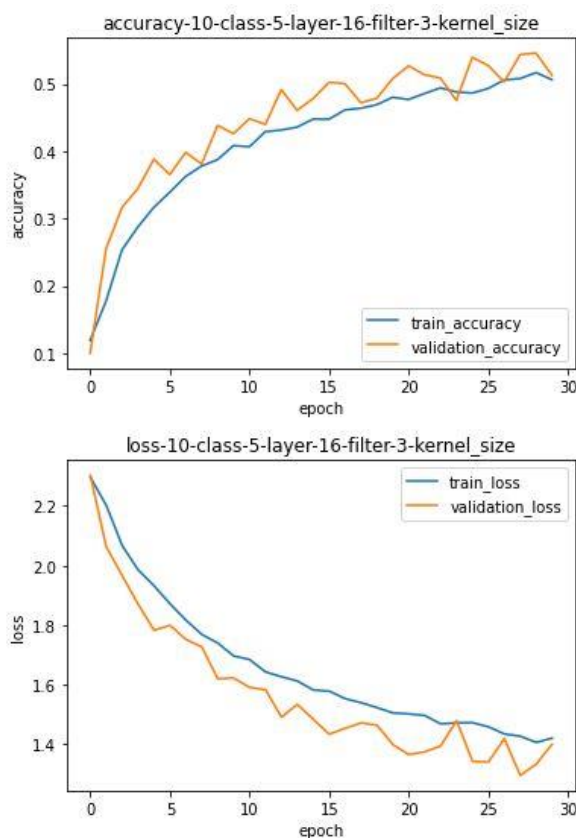


شکل ۶: نمودارهای **accuracy** و **loss** برای شبکه با ۷ جفت هر یک با ۳۲ کرنل به سائز ۳*۳

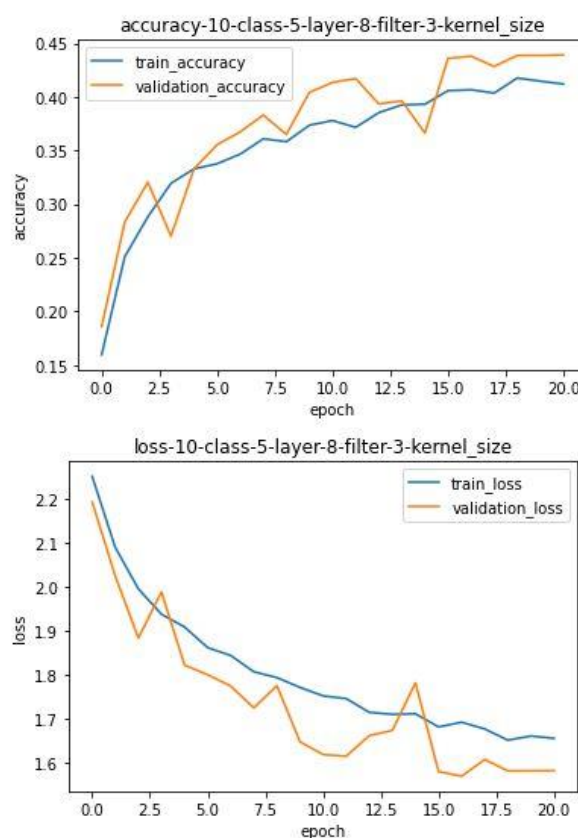
همان طور که مشاهده می‌کنید افزایش تعداد لایه‌ها تا ۵ جفت سبب بهبود دقت و هزینه می‌شود از جهتی دو نمودار مربوط به داده‌های تست و ارزیابی نیز با افزایش تعداد لایه‌ها تا ۵ جفت به یک دیگر نزدیک تر می‌شوند که نمایان گر بهبود **generalization** می‌باشد. اما افزایش تعداد لایه‌ها بیش از ۵ جفت سبب افت محسوس عملکرد شبکه می‌شود و همان طور که مشاهده می‌کنید نتایج برای ۶ و ۷ جفت مناسب نیست. دلیل آیین موضوع این است که با افزایش لایه‌ها ابعاد به دلیل استفاده از **pooling** کاهش می‌یابد و با کاهش بیش از اندازه این ابعاد سبب افت عملکرد در انتقال به بخش **classification** می‌شود.

تعداد کرنل‌های مختلف

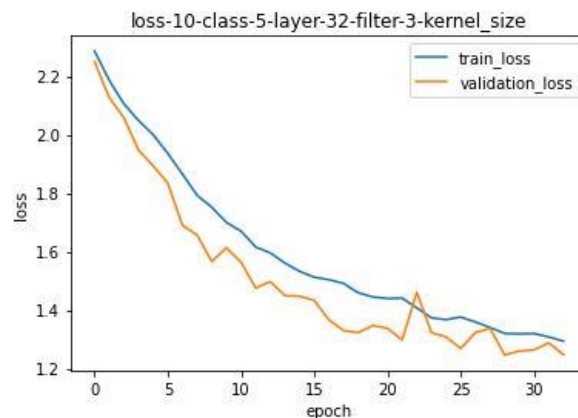
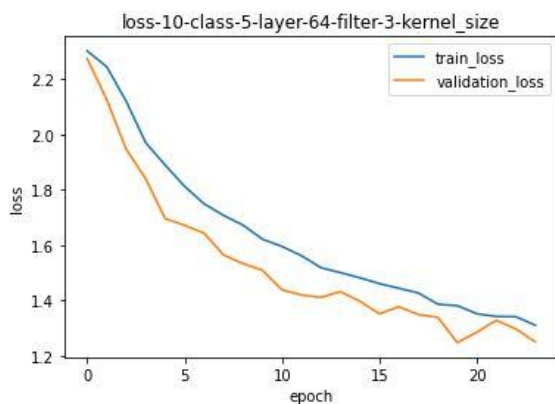
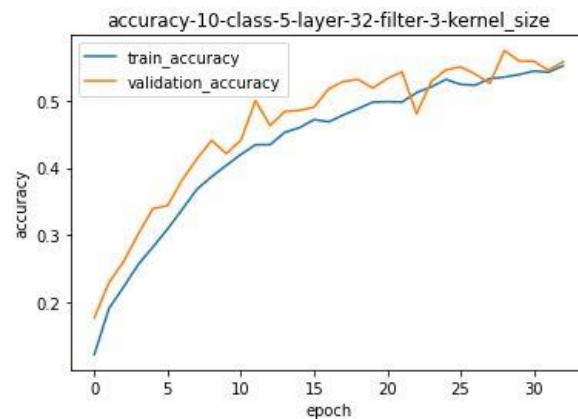
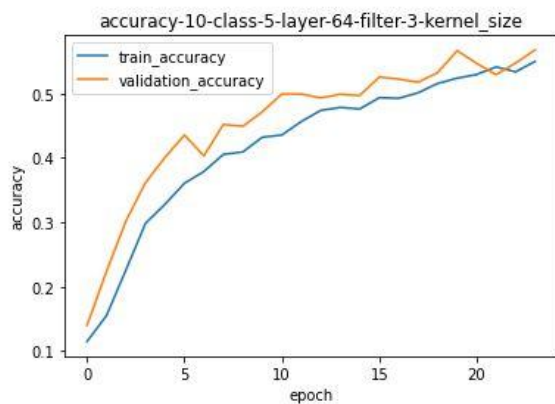
بهترین تعداد جفت در قسمت قبل ۵ شد. در این قسمت شبکه را با تعداد کرنل‌های متفاوت آموزش می‌دهیم. در زیر نمودارهای مربوط به **accuracy** و **loss** برای هر حالت به نمایش گذاشته شده است.



شکل ۸: نمودارهای **accuracy** و **loss** برای شبکه با ۵ جفت هر یک با ۱۶ کرنل به سایز ۳*۳

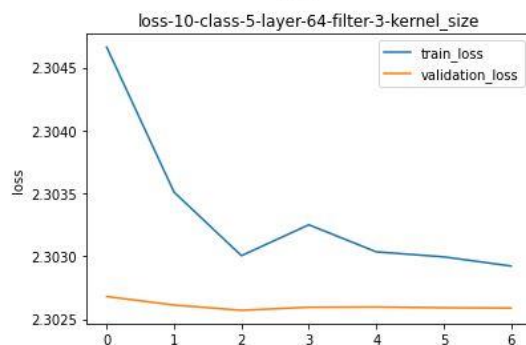
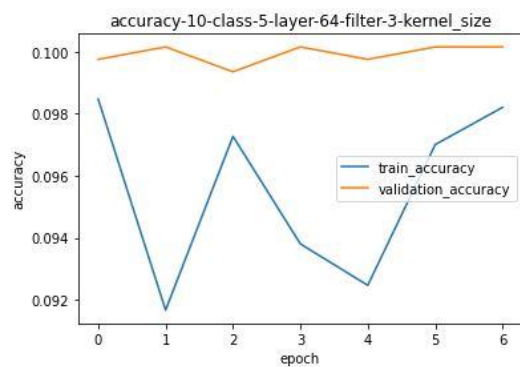


شکل ۷: نمودارهای **accuracy** و **loss** برای شبکه با ۵ جفت هر یک با ۸ کرنل به سایز ۳*۳



شکل ۱۰: نمودارهای **accuracy** و **loss** برای شبکه با ۵ جفت هر یک با ۶۴ کرنل به سایز ۳*۳

شکل ۹: نمودارهای **accuracy** و **loss** برای شبکه با ۵ جفت هر یک با ۳۲ کرنل به سایز ۳*۳

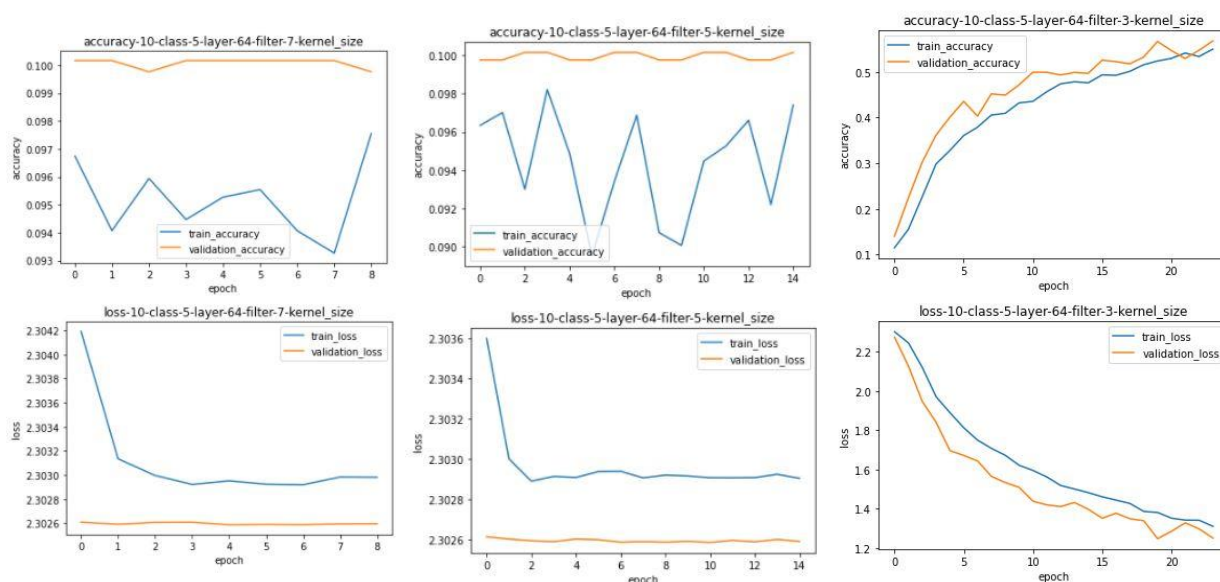


شکل ۱۱: نمودارهای **accuracy** و **loss** برای شبکه با ۵ جفت هر یک با ۱۲۸ کرنل به سایز ۳*۳

همان طور که مشاهده می‌کنید افزایش تعداد کرنل‌ها تا ۶۴ سبب بهبود دقت و هزینه می‌شود از جهتی دو نمودار مربوط به داده‌های تست و ارزیابی نیز با افزایش تعداد کرنل‌ها تا ۶۴ به یک دیگر نزدیک تر می‌شوند که نمایان گر بهبود **generalization** می‌باشد. اما افزایش تعداد کرنل‌ها بیش از ۶۴ سبب افت محسوس عملکرد شبکه می‌شود. دلیل این امر به دلیل افزایش بیش از اندازه پارامترهای شبکه باشد. درواقع می‌توان با استفاده بیشتر از **dropout** به عملکردی مشابه با ۶۴ کرنل در شبکه با ۱۲۸ کرنل رسید. در این آزمایش بهترین تعداد کرنل ۶۴ می‌باشد. این امکان وجود دارد تا تعداد کرنل‌ها را تا ۳۲۰۰۰ نیز افزایش دهیم اما این کار کمکی به بهبود عملکرد شبکه می‌کند.

اندازه‌های مختلف کرنل

در این قسمت مدل بهینه یافته شده از قسمت قبل را با ۵ جفت و هر یک با ۶۴ کرنل اما با اندازه‌های متفاوت مورد ارزیابی قرار می‌دهیم. در زیر نمودارهای مربوط به **accuracy** و **loss** برای هر حالت به نمایش گذاشته شده است.



شکل ۱۴: نمودارهای **accuracy** و **loss**

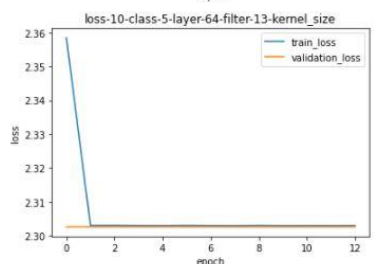
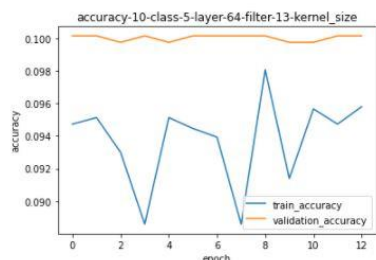
برای شبکه با ۵ جفت هر یک با ۱۲۸ کرنل به سائز ۷*۷

شکل ۱۳: نمودارهای **accuracy** و **loss** برای

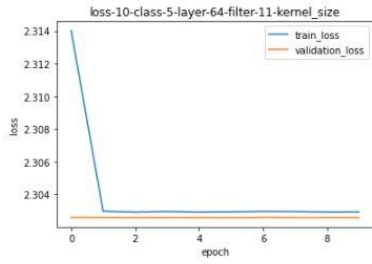
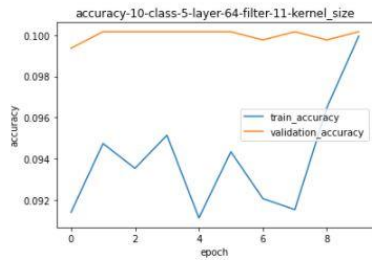
شبکه با ۵ جفت هر یک با ۱۲۸ کرنل به سائز ۵*۵

شکل ۱۲: نمودارهای **accuracy** و **loss** برای

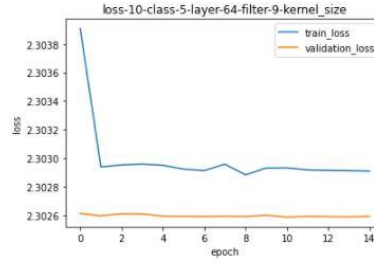
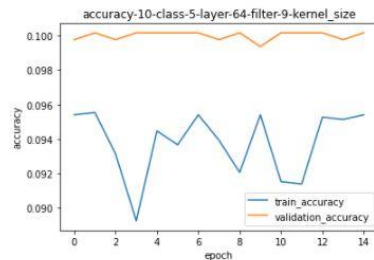
شبکه با ۵ جفت هر یک با ۱۲۸ کرنل به سائز ۳*۳



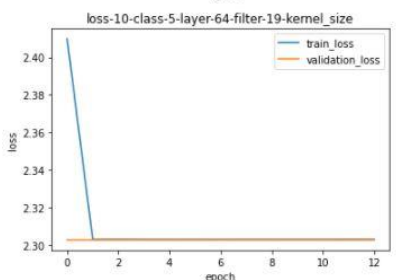
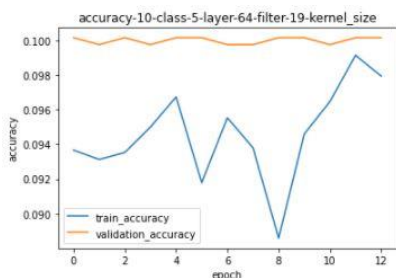
شکل ۱۷: نمودارهای accuracy و loss برای شبکه با ۵ جفت هر یک با ۱۲۸ کرنل به سائز ۱۳*۱۳



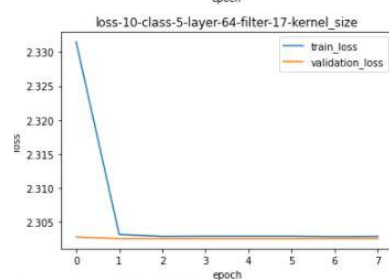
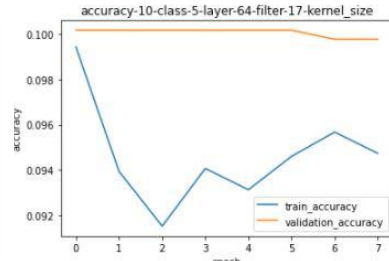
شکل ۱۶: نمودارهای accuracy و loss برای شبکه با ۵ جفت هر یک با ۱۲۸ کرنل به سائز ۱۱*۱۱



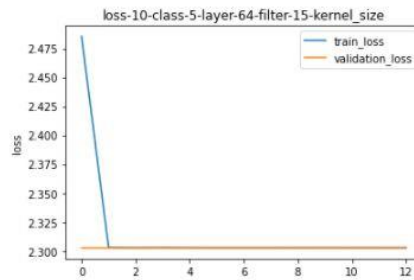
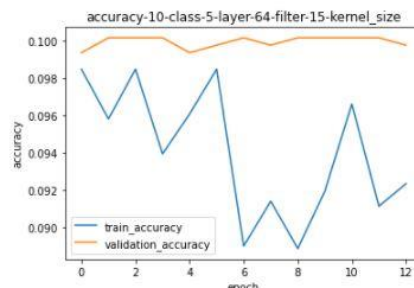
شکل ۱۵: نمودارهای accuracy و loss برای شبکه با ۵ جفت هر یک با ۱۲۸ کرنل به سائز ۹*۹



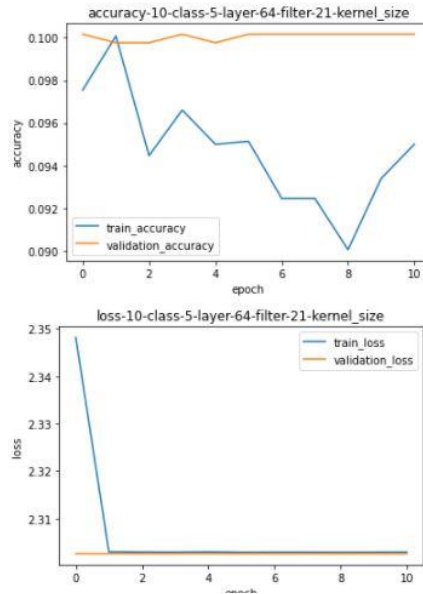
شکل ۲۰: نمودارهای accuracy و loss برای شبکه با ۵ جفت هر یک با ۱۲۸ کرنل به سائز ۱۹*۱۹



شکل ۱۹: نمودارهای accuracy و loss برای شبکه با ۵ جفت هر یک با ۱۲۸ کرنل به سائز ۱۷*۱۷



شکل ۱۸: نمودارهای accuracy و loss برای شبکه با ۵ جفت هر یک با ۱۲۸ کرنل به سائز ۱۵*۱۵



شکل ۲۱: نمودارهای accuracy و loss برای شبکه
با ۵ جفت هر یک با ۱۲۸ کرنل به سائز ۲۱*۲۱

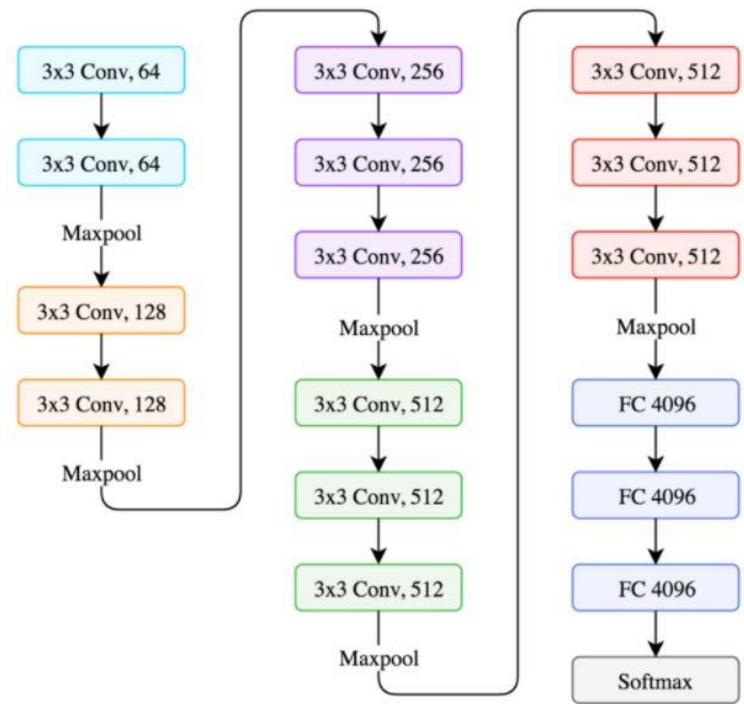
همان طور که انتظار داشتیم اندازه کرنل ۳*۳ بهتر از سایرین عمل می کند و افزایش اندازه کرنل سبب بهبود عملکرد نمی شود. در واقع می توان گفت استخراج ویژگی با اندازه کرنل پایین تر مناسب تر انجام می شود.

اندازه کرنل را می توان تا ۱۲۸*۱۲۸ افزایش داد که همان ابعاد تصویر ورودی به شبکه پس از resampling است. افزایش اندازه کرنل بیش از این مقدار مقدور نیست.

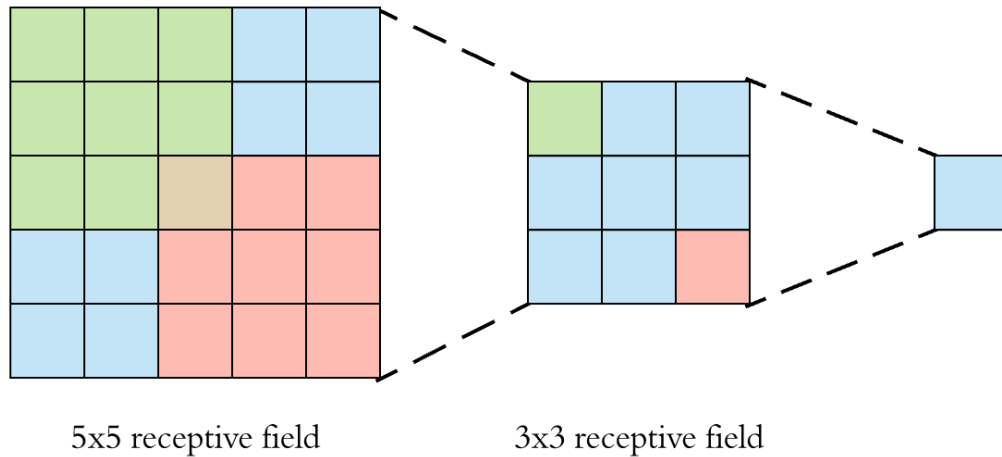
سوال دوم

این روش در مدل VGG پیاده شده است. در زیر شکل مدل VGG16 به نمایش گذاشته شده است. معماری شبکه به این صورت است که یک مجموعه ایی از لایه های کانولوشنی (stacked convolution) وجود دارد که پس از آن یک لایه maxpooling قرار می گیرد. شبکه تنها از کانولوشن های ۳*۳ در سراسر شبکه استفاده می کند. توجه داشته باشید که دو کانولوشن ۳*۳ پشت سر هم میدان دریافت موثر یک کانولوشن ۵*۵ را دارند. و سه کانولوشن ۳*۳ انباشته دارای میدان دریافتی یک ۷*۷ هستند. این موضوع برای حالت ۳*۳ و ۵*۵ به نمایش گذاشته شده است. در واقع هر صفحه کرنل یک ویژگی از تصویر استخراج می کند. کرنل های با ابعاد کوچک تر ویژگی های ساده تری را استخراج می کنند و هنگامی که چند لایه کانولوشنی را بدون زیر نمونه برداری با کرنل های کوچک استفاده می کنیم. ویژگی های استخراج شده با هم ادغام شده و عملکرد مناسب تری از یک صفحه کرنل با ابعاد بزرگ که ویژگی های پیچیده تر را استخراج میکند دارد. به بیان دیگر چند لایه کانولوشن کوچک موجب ادغام ویژگی های استخراج شده ساده همچون خطوط افقی، عمودی یا اوریب می شود اما یک لایه کانولوشنی به تنهایی سبب استخراج ویژگی های پیچیده تر به تنهایی می شود که مورد اول عملکرد بهتری در استخراج ویژگی ها دارد.

مزیت دیگر قرار دادن دو کانولوشن به جای یکی این است که از دو عملیات relu استفاده می کنیم و غیر خطی بودن بیشتر قدرت بیشتری به مدل می دهد.



شکل ۲۱: معماری شبکه VGG16



شکل ۲۱: مقایسه میدان موثر دو کرنل 3×3 و یک کرنل 5×5