

سوال اول

چند نمونه از تصاویر به همراه برچسب آن‌ها:

apple_pie



baby_back_ribs



baklava



beef_carpaccio



beef_tartare



beet_salad



beignets



bibimbap



bread_pudding



سوال دوم

اینکه تعداد زیادی لایه کانولوشنی با کرنل کوچک داشته باشیم تقریباً شبیه به ساختار شبکه VGG می‌شود. یکی از مزیت‌های این ساختار این است که وقتی این لایه‌های کانولوشن را اصطلاحاً پشت سر هم Stack می‌کنیم از تعداد توابع غیرخطی سازمانند Relu به تعداد دفعات بیشتری پشت سر هم استفاده می‌کنیم که باعث می‌شود قدرت مدل افزایش پیدا کند.

در حالتی که یک لایه کانولوشنی با اندازه کرنل بزرگ روی تصویر اعمال شود و نتیجه به یک لایه زیرنمونه‌برداری شود شبیه به ساختار شبکه AlexNet است که آموزش آن ۲ تا ۳ هفته زمان می‌برد و بسیار هزینه‌بر است. به همین دلیل بعد از آن ساختار VGG معرفی شد. در این نوع ساختار هزینه محاسباتی خیلی کمتر است و از طرفی خاصیت weight sharing باعث می‌شود که با وجود تعداد زیاد کرنل‌ها، تعداد وزن‌های BP کم باشد.

یک استدلال دیگر می‌تواند این باشد که کرنل‌های با سایز کوچک ویژگی‌های ساده و با درجه تجرید پایین را استخراج می‌کنند که ادغام این ویژگی‌ها می‌تواند مفیدتر از یک ویژگی با درجه تجرید بالا که توسط یک کرنل با سایز بزرگ استخراج شده است عمل کند.

سوال سوم

قسمت اول: بررسی تعداد جفت لایه‌های کانولوشن زیرنمونه‌برداری

نتایج بررسی برای ۱۰۱ کلاس در جدول صفحه شماره ۱ آورده شده است. در تمام آزمایش‌ها تعداد epochها برابر ۱۰ در نظر گرفته شده است. همچنین تعداد فیلترها برابر ۱۲ و سایز فیلترها برابر ۳ در نظر گرفته شده است.

نوع پدینگ: در حالت Same به دلیل حاشیه ای که در تصاویر اضافه میشود، اطلاعات حفظ میشوند و سرعت کاهش سایز تصاویر بسیار کمتر از حالت Valid است. بنابراین اگر از حالت Valid استفاده کنیم تعداد لایه‌ها محدود میشود اما این محدودیت در حالت Same وجود ندارد و در حالت Same سرعت کاهش سایز تصاویر کمتر است. بنابراین میتوان تعداد لایه‌های بیشتری برای شبکه در نظر گرفت و شبکه عمیق تری طراحی کرد.

همانطور که مشاهده می‌شود دقت روی داده‌های تست برای ۴ جفت لایه کانولوشن زیرنمونه‌برداری بهترین است. روند دقت ابتدا افزایشی و سپس کاهشی است.

می‌توان علت را اینگونه بیان کرد که وقتی تعداد لایه‌ها زیاد می‌شود از آنجایی که مرتباً ابعاد تصویر در حال کاهش است از یک جایی به بعد اطلاعاتی از تصویر که به لایه آخر برای دسته‌بندی می‌رسد خیلی کم هستند و دقت

مدل پایین می‌آید. از طرفی وقتی تعداد لایه‌ها خیلی کم است اطلاعات و فیچرهای زیادی به لایه آخر که مربوط به دسته‌بندی است منتقل می‌شوند که موجب می‌شوند دسته‌بندی به خوبی صورت نگیرد.

تعداد این جفت لایه‌ها را فقط تا عدد ۷ می‌توانیم جلو ببریم چون ابعاد تصویر 128×128 است و از آن جایی که padding را برای لایه کانولوشن "same" داده‌ایم و لایه زیرنمونه‌برداری 2×2 استفاده کرده‌ایم، ابعاد تصویر بعد از هر جفت لایه نصف می‌شود. بنابراین نمی‌توان تعداد لایه‌ها را تا ۵۰۰ افزایش داد زیرا ابعاد تصویر اجازه نمی‌دهد و تصویر نابود می‌شود.

قسمت دوم: بررسی تعداد کرنل‌های کانولوشن

در این قسمت با ۴ جفت لایه که بیشترین دقت را از قسمت قبل به دست آورده بود تعداد کرنل‌های مختلف را بررسی می‌کنیم. نتایج این بررسی در جدول شماره ۲ آورده شده است. همانطور که مشاهده می‌شود تعداد کرنل ۶۴ بیشترین دقت را دارد که برابر ۲۵ درصد است.

افزایش تعداد کرنل‌ها تا یک جایی باعث افزایش دقت می‌شود اما از آن به بعد مدل دچار اورفیت شده و با اینکه loss کم می‌شود اما دقت روی داده‌های تست پایین می‌آید. علتش این است که افزایش تعداد کرنل‌ها باعث می‌شود تعداد پارامترها و در نتیجه میزان محاسبات افزایش یابد و شبکه دچار اورفیت می‌شود.

برای اطمینان از این امر، روی شبکه‌ای که ۱۲۸ تا کرنل دارد تعداد Dropout‌های بیشتری اعمال شد و مشاهده شد که دقت (در حالت ۱۰ کلاس) به ۵۷ درصد رسید. بنابراین علت کاهش دقت مربوط به اورفیت بوده است.

تعداد کرنل‌ها را نمی‌توانیم تا ۳۲۰۰۰ زیاد کنیم زیرا کرنل‌ها در واقع آرایه‌های دو بعدی هستند و از یک جایی به بعد برای تعریف تعداد زیاد کرنل با کمبود حافظه مواجه می‌شویم.

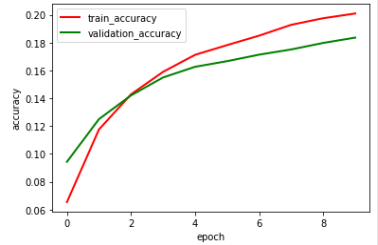
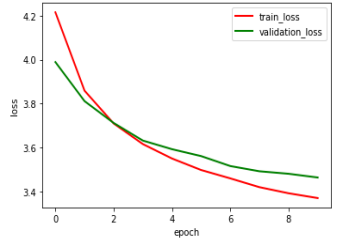
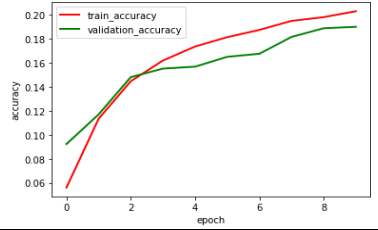
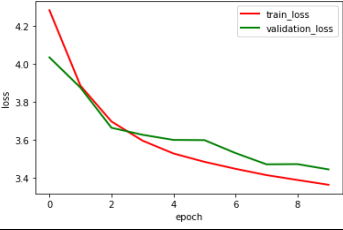
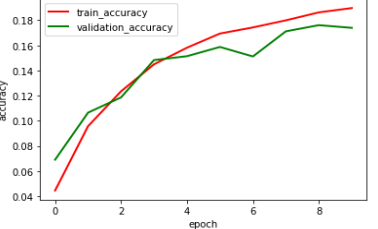
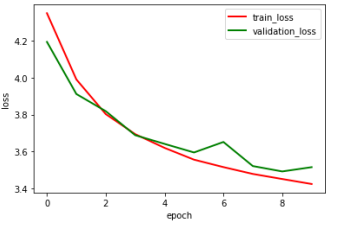
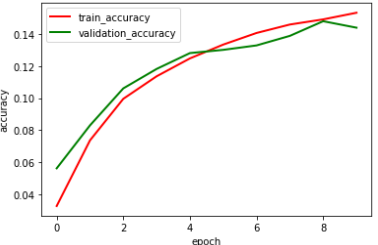
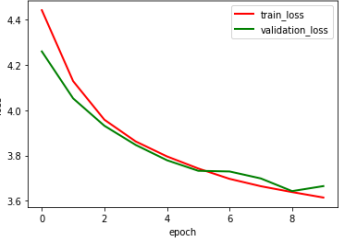
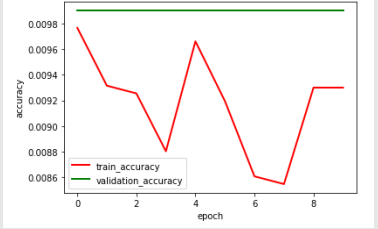
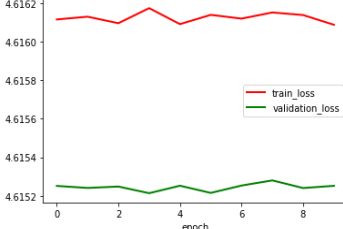
قسمت سوم: بررسی سایز کرنل‌های کانولوشن

در این قسمت با ۴ جفت لایه که بیشترین دقت را از قسمت اول به دست آورده بود و تعداد کرنل‌های ۶۴ که بیشترین دقت را از قسمت دوم به دست آورده بود تاثیر سایز کرنل را بررسی می‌کنیم. نتایج این بررسی در جدول شماره ۳ آورده شده است. همانطور که مشاهده می‌شود سایز کرنل ۳ بیشترین دقت را دارند که برابر ۲۵ درصد و هر چقدر سایز کرنل‌ها بیشتر می‌شود دقت پایین تر می‌آید.

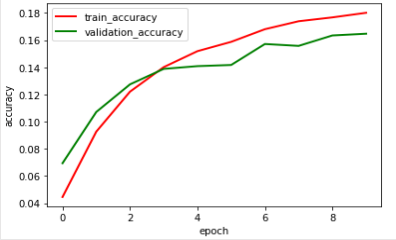
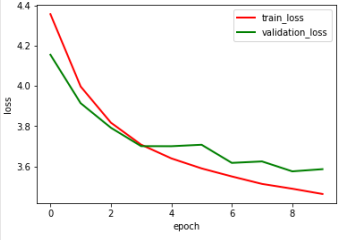
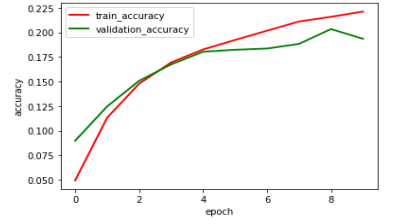
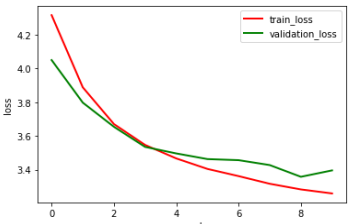
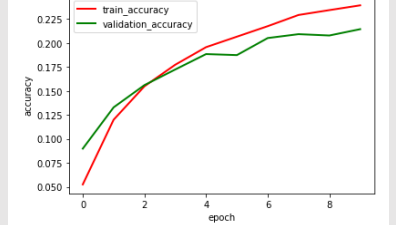
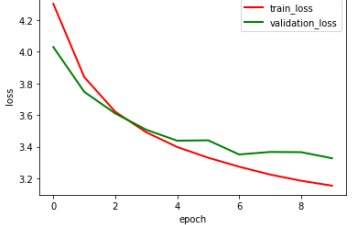
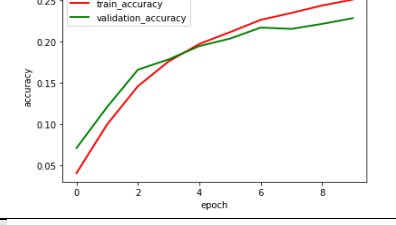
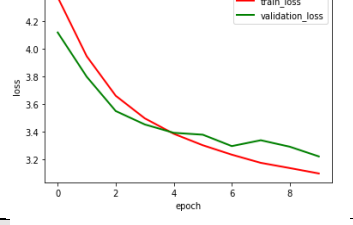
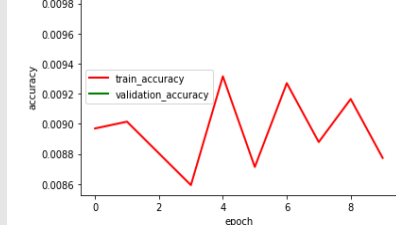
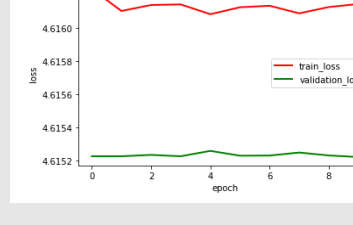
اندازه کرنل‌های بزرگتر باعث می‌شود تعداد پارامترها و در نتیجه میزان محاسبات افزایش یابد. این امر می‌تواند موجب اورفیت شدن مدل شود.

از طرفی علت پایین آمدن دقت با افزایش سایز کرنل‌ها را می‌توان با همان استدلالی که در سوال ۲ آورده شد بیان کرد. همانطور که گفتیم کرنل‌های با سایز پایینتر ویژگی‌های ساده و با درجه تجرید پایین را استخراج می‌کنند که ادغام این ویژگی‌ها می‌تواند مفیدتر از یک ویژگی با درجه تجرید بالا که توسط یک کرنل با سایز بزرگ استخراج شده است عمل کند.

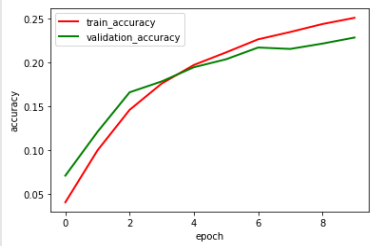
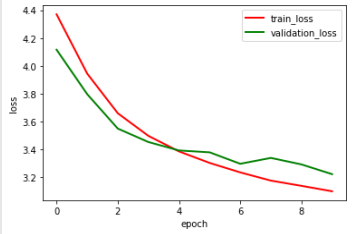
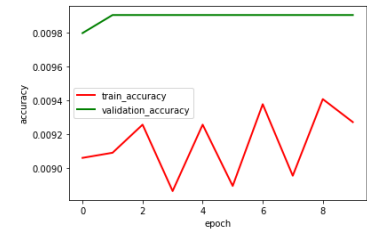
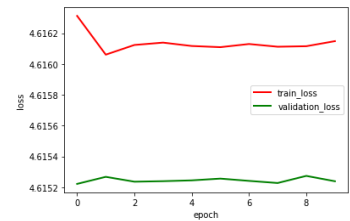
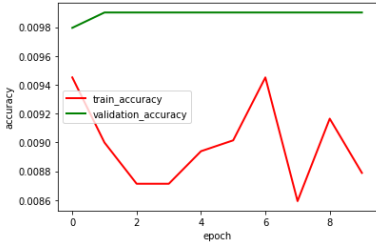
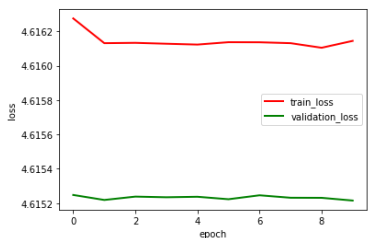
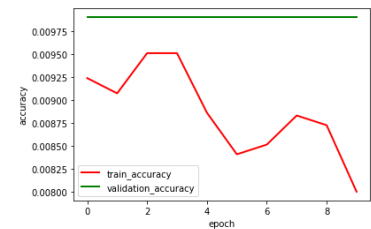
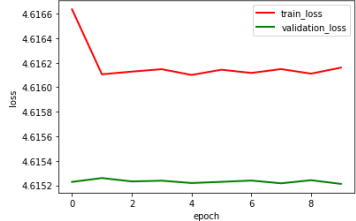
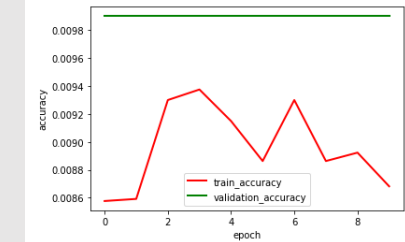
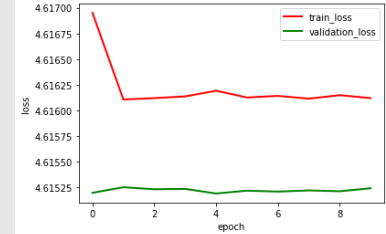
سایز کرنل‌ها را نمی‌توانیم تا ۲۰۱ زیاد کنیم زیرا ابعاد تصویری که ما در نظر گرفته‌ایم $128 * 128$ است و سایز کرنل نمی‌تواند از ۱۲۸ بیشتر باشد.

تعداد لایه‌ها	دقت (تست)	Loss (تست)	نمودار accuracy	نمودار Loss
3	0.19	3.390		
4	0.21	3.282		
5	0.20	3.327		
6	0.16	3.5326		
7	0.0099	4.615		

جدول 1 بررسی تعداد لایه‌ها برای ۱۰۱ کلاس

تعداد کرنل‌ها	دقت (تست)	Loss (تست)	نمودار accuracy	نمودار Loss
8	0.18	3.43		
16	0.22	3.19		
32	0.23	3.17		
64	0.25	3.05		
128	0.0099	4.61		

جدول 2 بررسی تعداد کرنل‌ها برای ۱۰۱ کلاس

سایز کرنل‌ها	دقت (تست)	Loss (تست)	نمودار accuracy	نمودار Loss
3	0.25	3.05		
5	0.0099	4.61		
7	0.0099	4.61		
9	0.0099	4.61		
11	0.0099	4.61		

جدول 3 بررسی سایز کرنل‌ها برای ۱۰۱ کلاس

در ادامه به دلیل اینکه دقت‌ها برای ۱۰۱ کلاس تا حد زیادی پایین بود و اینکه افزایش تعداد epoch ها برای ۱۰۱ کلاس زمان زیادی را برای اجرا طلب می‌کرد آزمایش‌های سوال ۳ برای ۱۰ کلاس که به صورت رندوم از کلاس‌های داده‌ها انتخاب شده‌اند انجام گرفته است. ۱۰ کلاس انتخاب شده این موارد هستند:

['beef_tartare', 'chicken_curry', 'chocolate_mousse', 'french_toast', 'fried_rice', 'hot_dog', 'lasagna', 'oysters', 'pizza', 'takoyaki']

قسمت اول: بررسی تعداد جفت لایه‌های کانولوشن زیرنمونه‌برداری

نتایج بررسی برای ۱۰ کلاس در جدول شماره ۴ آورده شده است. در تمام آزمایش‌ها تعداد epoch ها برابر ۲۰ در نظر گرفته شده است. هم‌چنین تعداد فیلترها برابر ۱۲ و سائز فیلترها برابر ۳ در نظر گرفته شده است.

همانطور که مشاهده می‌شود دقت روی داده‌های تست برای ۵ جفت لایه کانولوشن زیرنمونه‌برداری بهترین و برابر ۵۰ درصد است. روند دقت ابتدا افزایشی و سپس کاهشی است.

می‌توان علت را اینگونه بیان کرد که وقتی تعداد لایه‌ها زیاد می‌شود از آنجایی که مرتباً ابعاد تصویر در حال کاهش است از یک جایی به بعد اطلاعاتی از تصویر که به لایه آخر برای دسته‌بندی می‌رسد خیلی کم هستند و دقت مدل پایین می‌آید. از طرفی وقتی تعداد لایه‌ها خیلی کم است اطلاعات و فیچرهای زیادی به لایه آخر که مربوط به دسته‌بندی است منتقل می‌شوند که موجب می‌شوند دسته‌بندی به خوبی صورت نگیرد.

تعداد این جفت لایه‌ها را فقط تا عدد ۷ می‌توانیم جلو ببریم چون ابعاد تصویر 128×128 است و از آن جایی که padding را برای لایه کانولوشن "same" داده‌ایم و لایه زیرنمونه‌برداری 2×2 استفاده کرده‌ایم، ابعاد تصویر بعد از هر جفت لایه نصف می‌شود. بنابراین نمی‌توان تعداد لایه‌ها را تا ۵۰۰ افزایش داد زیرا ابعاد تصویر اجازه نمی‌دهد.

در شکل زیر ساختار شبکه برای ۳ جفت لایه‌ی کانولوشنی - زیرنمونه‌برداری نشان داده شده است:

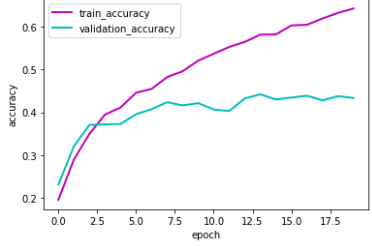
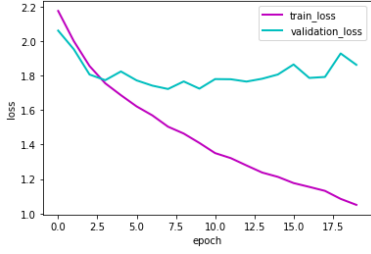
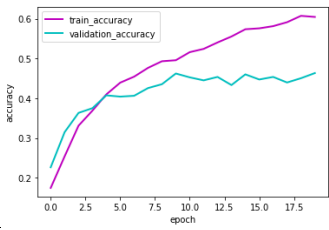
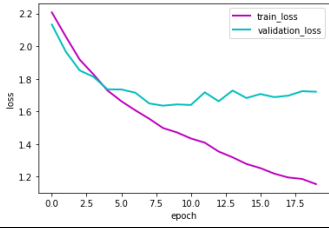
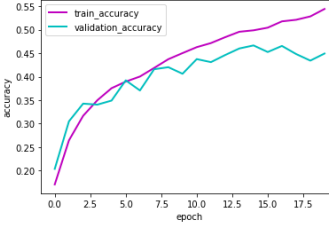
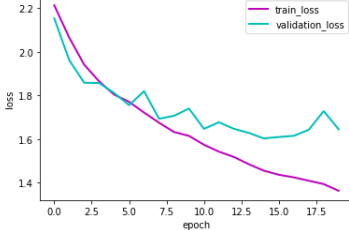
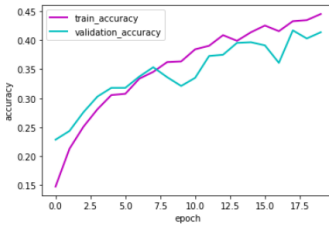
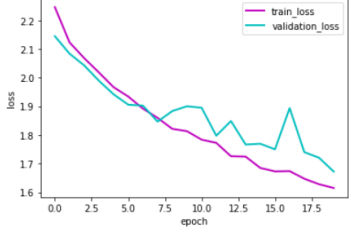
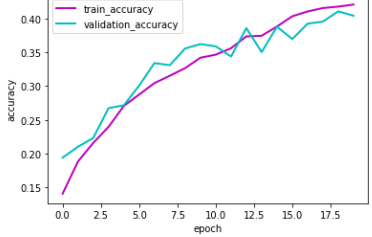
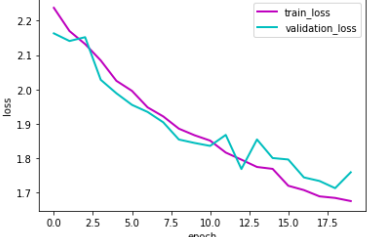
Layer (type)	Output Shape	Param #
inp (InputLayer)	[(None, 128, 128, 3)]	0
conv2d (Conv2D)	(None, 128, 128, 12)	336
max_pooling2d (MaxPooling2D)	(None, 64, 64, 12)	0
conv2d_1 (Conv2D)	(None, 64, 64, 12)	1308
max_pooling2d_1 (MaxPooling2D)	(None, 32, 32, 12)	0
conv2d_2 (Conv2D)	(None, 32, 32, 12)	1308
max_pooling2d_2 (MaxPooling2D)	(None, 16, 16, 12)	0
flatten (Flatten)	(None, 3072)	0
dense (Dense)	(None, 128)	393344
dense_1 (Dense)	(None, 10)	1290
Total params: 397,586		
Trainable params: 397,586		
Non-trainable params: 0		

قسمت دوم: بررسی تعداد کرنل‌های کانولوشن

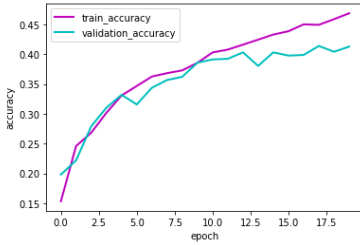
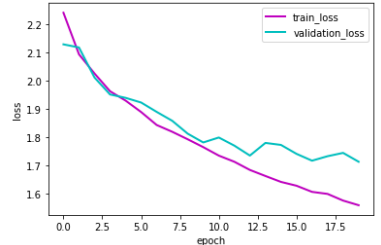
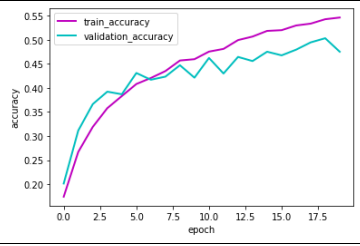
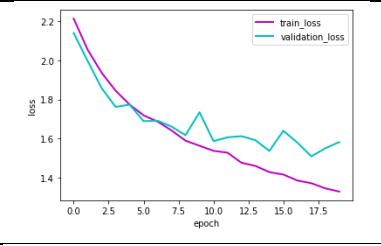
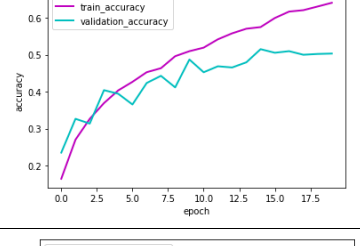
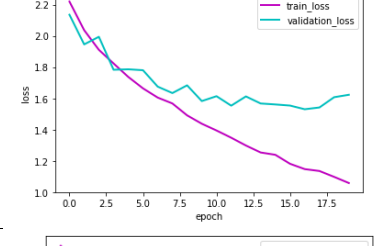
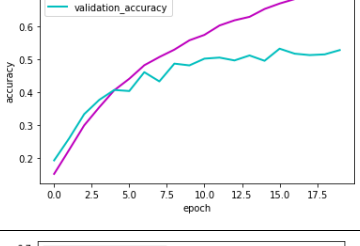
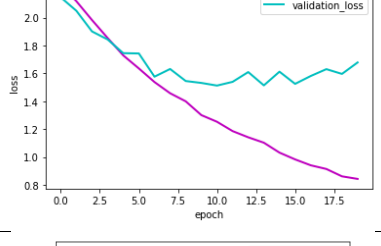
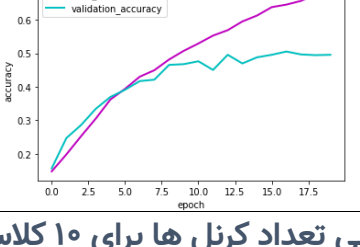
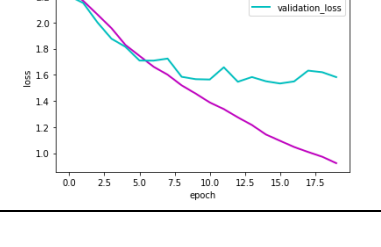
در این قسمت با ۵ جفت لایه که بیشترین دقت را از قسمت قبل به دست آورده بود تعداد کرنل‌های مختلف را بررسی می‌کنیم. نتایج این بررسی در جدول شماره ۵ آورده شده است. همانطور که مشاهده می‌شود تعداد کرنل ۳۲ و ۶۴ بیشترین دقت را دارند که برابر ۵۵ درصد است اما با توجه به نمودار loss و accuracy می‌توان مشاهده کرد که برای تعداد ۶۴ کرنل اورفیت داریم به همین دلیل ۳۲ کرنل را به عنوان بهترین انتخاب می‌کنیم.

همانطور که مشاهده می‌شود افزایش تعداد کرنل‌ها تا یک جایی باعث افزایش دقت می‌شود اما از آن به بعد مدل دچار اورفیت شده و با اینکه loss کم می‌شود اما دقت روی داده‌های تست پایین می‌آید. برای اطمینان از این امر، روی شبکه ای که ۱۲۸ تا کرنل دارد تعداد Dropout‌های بیشتری اعمال شد و مشاهده شد که دقت (در حالت ۱۰ کلاسه) به ۵۷ درصد رسید. بنابراین علت کاهش دقت مربوط به اورفیت بوده است.

تعداد کرنل‌ها را نمی‌توانیم تا ۳۲۰۰۰ زیاد کنیم زیرا کرنل‌ها در واقع آرایه‌های دو بعدی هستند و از یک جایی به بعد برای تعریف تعداد زیاد کرنل با کمبود حافظه مواجه می‌شویم.

تعداد لایه‌ها	دقت (تست)	Loss (تست)	نمودار accuracy	نمودار Loss
3	0.47	1.59		
4	0.49	1.52		
5	0.50	1.46		
6	0.45	1.58		
7	0.41	1.66		

جدول 4 بررسی تعداد لایه‌ها برای ۱۰ کلاس

تعداد کرنل‌ها	دقت (تست)	Loss (تست)	نمودار accuracy	نمودار Loss
8	0.46	1.57		
16	0.51	1.44		
32	0.55	1.42		
64	0.55	1.35		
128	0.54	1.40		

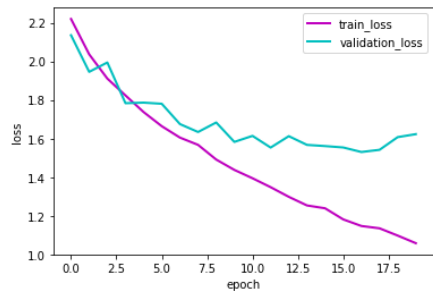
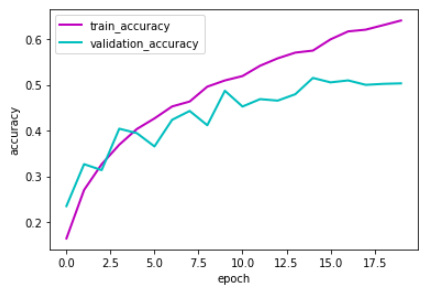
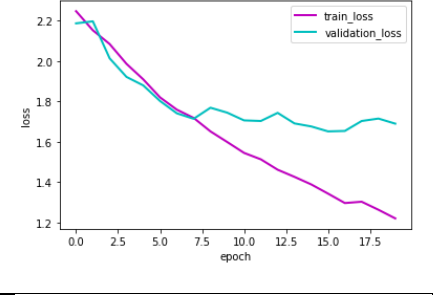
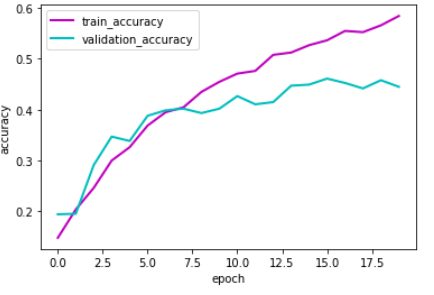
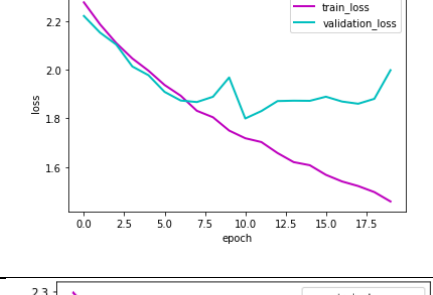
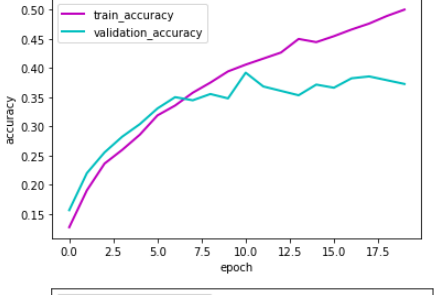
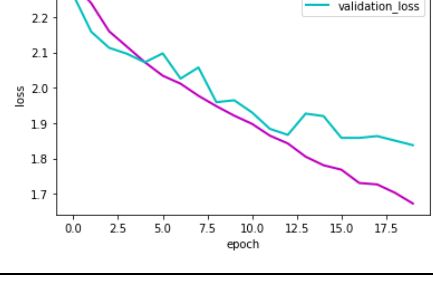
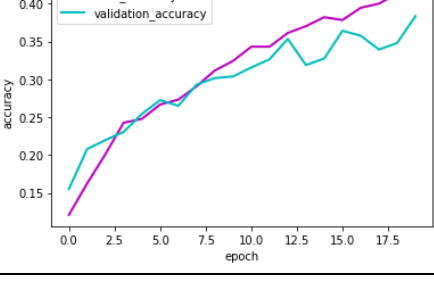
جدول 5 بررسی تعداد کرنل ها برای ۱۰ کلاس

قسمت سوم: بررسی سایز کرنل‌های کانولوشن

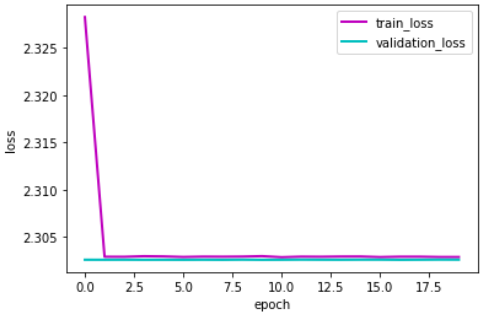
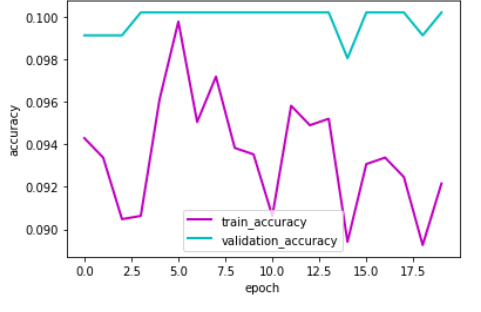
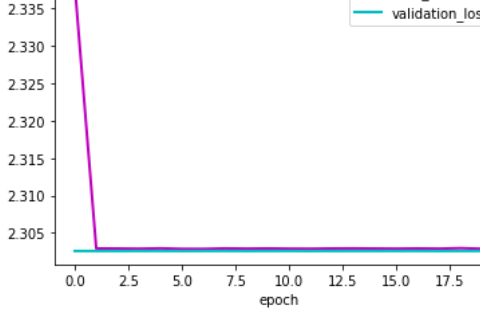
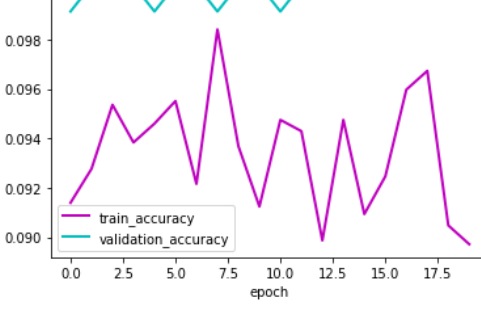
در این قسمت با ۵ جفت لایه که بیشترین دقت را از قسمت اول به دست آورده بود و تعداد کرنل‌های ۳۲ که بیشترین دقت را از قسمت دوم به دست آورده بود تاثیر سایز کرنل را بررسی می‌کنیم. نتایج این بررسی در جدول شماره ۶ آورده شده است. همانطور که مشاهده می‌شود سایز کرنل ۳ بیشترین دقت را دارند که برابر ۵۵ درصد و هر چقدر سایز کرنل‌ها بیشتر می‌شود دقت پایین تر می‌آید.

اندازه کرنل‌های بزرگتر باعث می‌شود تعداد پارامترها و در نتیجه میزان محاسبات افزایش یابد. این امر می‌تواند موجب اورفیت شدن مدل شود. به وضوح در نمودارهای loss و accuracy این مورد قابل مشاهده است.

سایز کرنل‌ها را نمی‌توانیم تا ۲۰۱ زیاد کنیم زیرا ابعاد تصویری که ما در نظر گرفته‌ایم 128×128 است و سایز کرنل نمی‌تواند از ۱۲۸ بیشتر باشد.

نمودار Loss	نمودار accuracy	Loss (تست)	دقت (تست)	سایز کرنل‌ها
		1.42	0.55	3
		1.55	0.49	5
		1.74	0.40	7
		1.80	0.39	9

		1.98	0.29	11
		2.30	0.10	13
		2.30	0.10	15
		2.30	0.10	17

		2.30	0.10	19
		2.30	0.10	21

جدول 6 بررسی سائز کرنل‌ها برای ۱۰ کلاس