



دانشگاه صنعتی امیرکبیر  
(پلی تکنیک تهران)  
دانشکده مهندسی کامپیوتر

## **پروژه چهارم درس رایانش عصبی**

نگارش

سیدمهدی میرفندرسکی

مدرس

دکتر رضا صفابخش

آذر ۱۴۰۱

## فهرست مطالب

۲	۱ تئوری - سوال اول
۲	۲ تئوری - سوال دوم
۲	۳ تئوری - سوال سوم
۲	۴ پیاده سازی - سوال اول
۸	۵ پیاده سازی - سوال سوم
۱۰	۶ پیاده سازی - سوال چهارم
۱۲	۷ پیاده سازی - سوال پنجم
۱۲	۸ پیاده سازی - سوال ششم

## ۱ تئوری - سوال اول

با توجه به فلسفه وجودی این نوع شبکه‌ها، جهت جلوگیری از کاهش تاثیر تغییر وزن‌ها، هرچه شبکه عمیق‌تر باشد، بهتر است طول این اتصالات بیشتر باشد. اگر فرض کنیم یک شبکه ۱۰۰ لایه پنهان دارد، طبیعتاً یک اتصال باقیماندگی از لایه سوم به لایه اول تأثیری چندانی نخواهد داشت، پس بهتر است طول این اتصال بیشتر باشد تا تأثیر تغییر وزن افزایش یابد. از طرفی با استدلال مشابه، هرچه شبکه عمق کمتری داشته باشد، طول اتصالات بهتر است کاهش یابد.

## ۲ تئوری - سوال دوم

Average pooling می‌تواند قدرت کلی یک ویژگی را با عبور گرادیان‌ها از میان همه مقادیر بهتر نشان دهد. از طرفی در max pooling گرادیان تنها از بیشترین مقدار عبور می‌کند. با این استدلال در DenseNet، Average pooling مناسب‌تر است. به بیان دیگر با ماکزیمم‌گیری امکان از بین رفتن ویژگی‌ها وجود دارد اما در میانگین‌گیری همه پیکسل‌ها تأثیر داده می‌شوند.

## ۳ تئوری - سوال سوم

به نظر من، مسئله دسته‌بند را شبیه تخمین تابع در نظر بگیریم و همچنین می‌دانیم هر مسئله تخمین تابع را می‌توان با یک شبکه دارای یک لایه پنهان تخمین زد. بدین ترتیب در سمت دسته‌بند یک ابرپارامتر خواهیم داشت. همچنین در قسمت استخراج ویژگی‌ها چندین ابرپارامتر وجود دارند. بنابراین چندین مقدار مانند تمرین دوم برای هر ابرپارامتر در نظر گرفته و تعدادی شبکه با ترکیبات مختلف آن‌ها آموزش می‌دهیم. و در نهایت بهترین را گزارش می‌دهیم. به ابرپارامترهای در نظر گرفته شده در سوال بعدی اشاره می‌شود. البته در قسمت استخراج می‌توان چند نکته را در نظر گرفت. با توجه به نوع تصاویر می‌توان اندازه فیلتر را بهینه انتخاب کرد (هرچه کوچک‌تر جزئیات بیش‌تر و بالعکس). موارد مشابه نیز برای تعداد فیلتر (عمق تصاویر) را می‌توان در نظر گرفت.

## ۴ پیاده‌سازی - سوال اول

در این قسمت برای آزمون و خطای شبکه‌ها ۹۷ شبکه آموزش داده شد. درواقع با در نظر گرفتن مقادیر زیر برای تعدادی ابرپارامتر نتایجی تولید شد که به ضمیمه ارسال می‌شود. ابرپارامترها و مقادیر مختلف آن‌ها در جدول زیر مشاهده می‌شود.

مقادیر	ابرپارامتر
۱۶ ۳۲ ۶۴ ۱۲۸	kernel numbers
۳ ۵ ۷	kernel size
۱ ۲	strides
same	padding
relu	activation
avg, max	pooling app
۲	pool size
۱۲۸ ۲۵۶	hidden layer
False	dropout enable

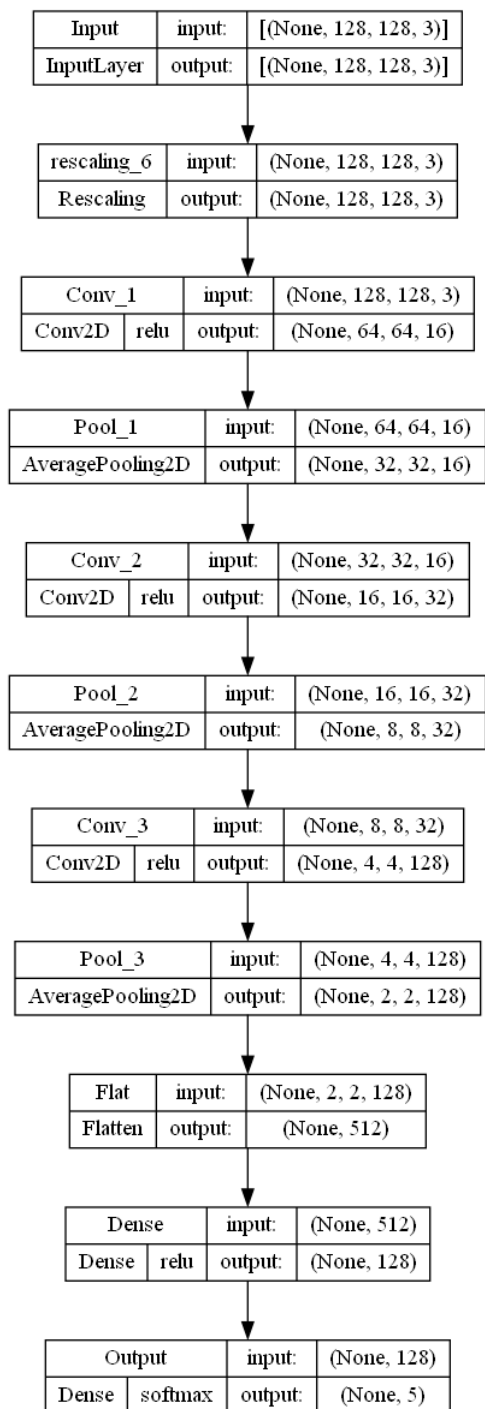
همچنین تمام شبکه‌ها با کالک زیر آموزش داده شدند.

es-callback = EarlyStopping(monitor="val-loss", patience=6, restore-best-weights=True)

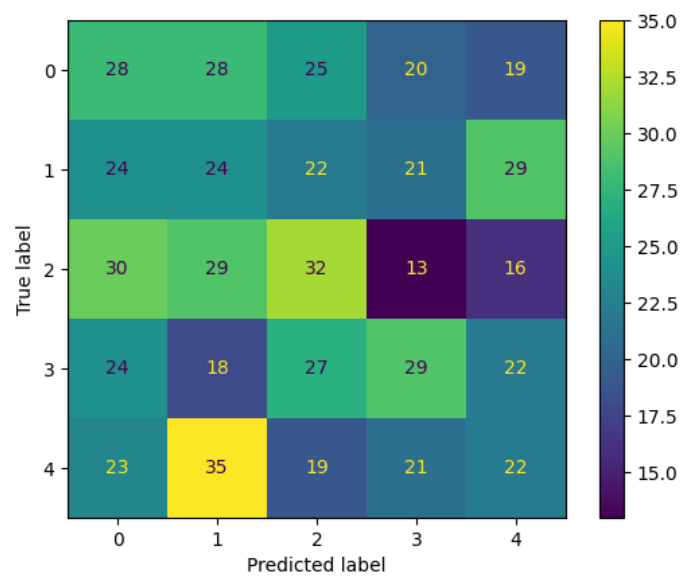
شبکه‌ها دوبار آموزش داده شدند (دو فایل ضمیمه شده). ابرپارامترهای دو شبکه که بهترین مقادیر صحت برای مجموعه اعتبارسنجی برای هربار اجرا را دارند، به شرح زیر می‌باشند.

ابرپارامتر	مدل ۱	مدل ۲
kernel numbers	۱۶ ۳۲ ۱۲۸	۱۶ ۶۴ ۱۲۸
kernel size	۵	۳
strides	۲	۲
padding	same	same
activation	relu	relu
pooling app	avg	max
pool size	۲	۲
hidden layer	۱۲۸	۲۵۶
dropout enable	False	False

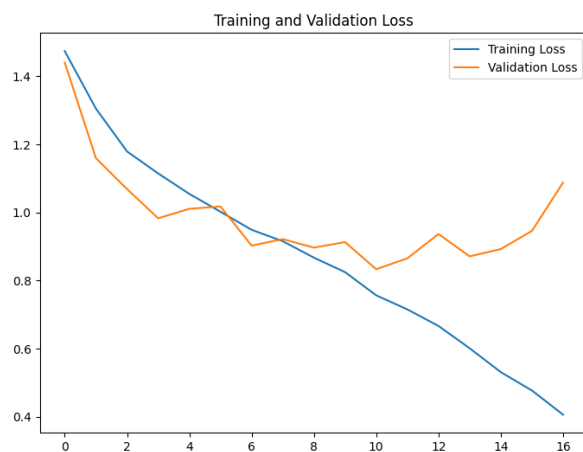
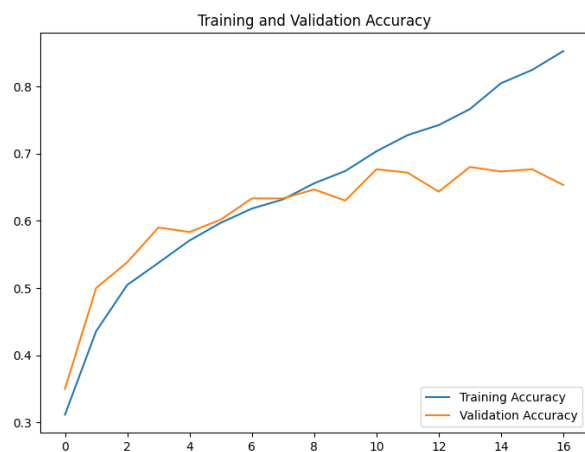
برای مدل اول گراف مصور مدل، ماتریس درهم‌ریختگی، نمودارهای خطا و صحت به ترتیب آورده شده‌اند.



شکل ۱: گراف مصور مدل

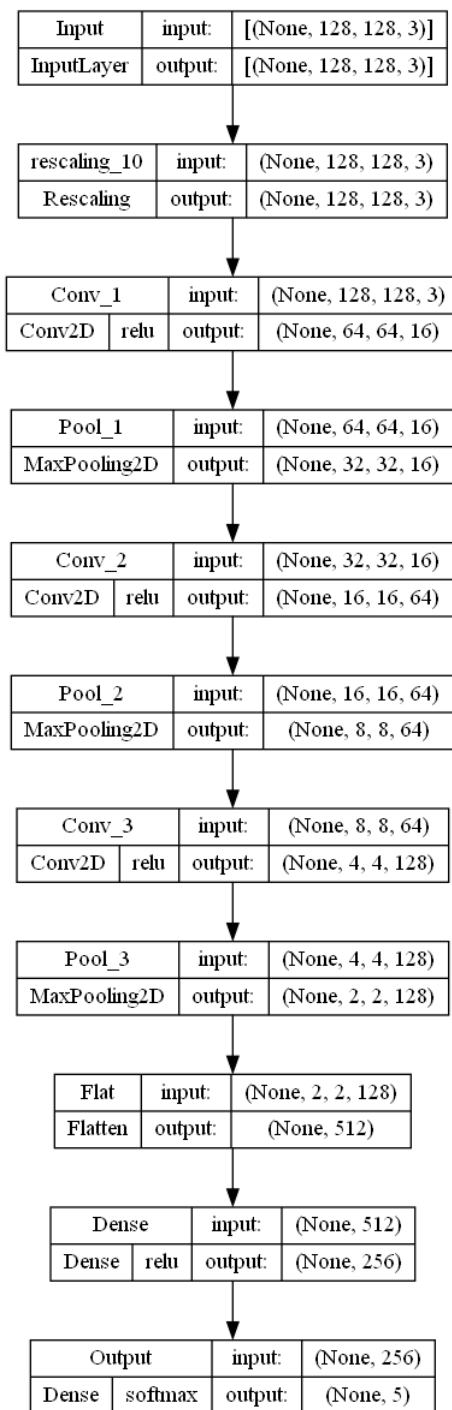


شکل ۲: ماتریس درهم‌ریختگی

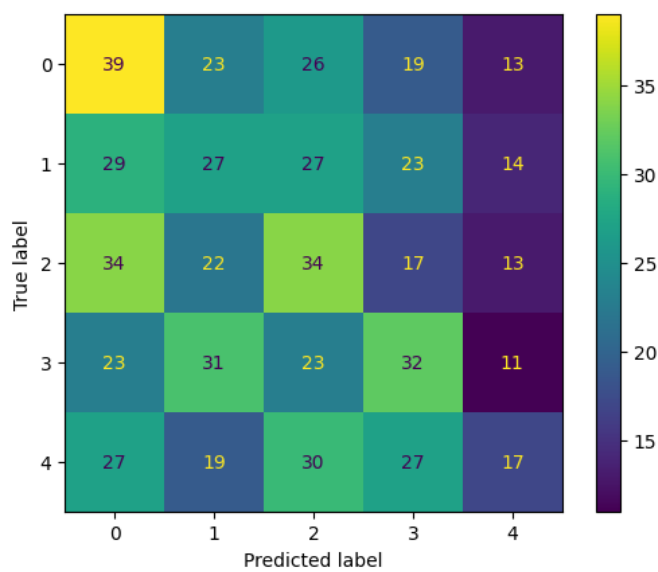


شکل ۳: نمودارهای خطا و صحت

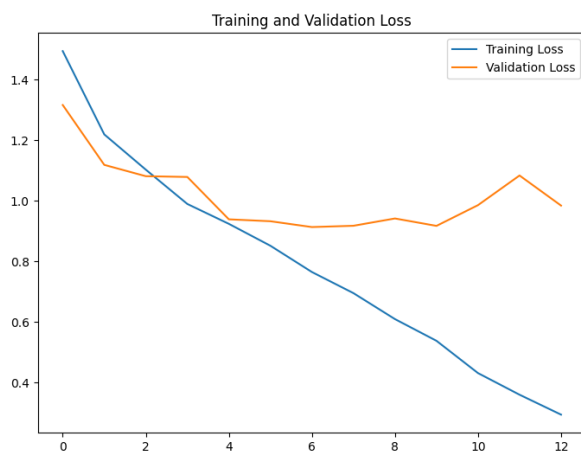
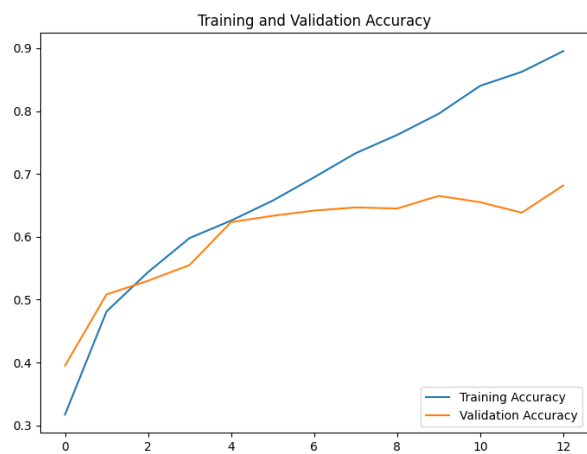
همانطور که قابل مشاهده است، مقدار ای پاک ۱۳ مقدار قابل قبولی خواهد بود. بعد از آن مدل دچار بیش‌برازش می‌شود. برای مدل دوم گراف مصور مدل، ماتریس درهم‌ریختگی، نمودارهای خطا و صحت به ترتیب آورده شده‌اند.



شکل ۴: گراف مصور مدل



شکل ۵: ماتریس درهم‌ریختگی



شکل ۶: نمودارهای خطا و صحت



همانطور که قابل مشاهده است، مقدار ای پاک ۶ تا ۹ مقدار قابل قبولی خواهد بود. بعد از آن مدل دچار بیش‌برازش می‌شود.

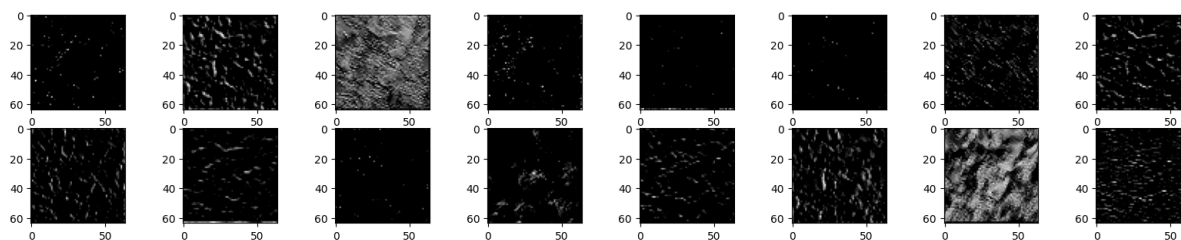
## ۵ پیاده‌سازی - سوال سوم

برای این بخش ابتدا از یکی از مدل‌های آموزش داده شده استفاده می‌کنیم. و لایه اول کانولوشنی به عنوان خروجی در نظر گرفته می‌شود. همچنین سه عکس از مجموعه آموزشی در نظر گرفته می‌شود که در ادامه آورده شده است.

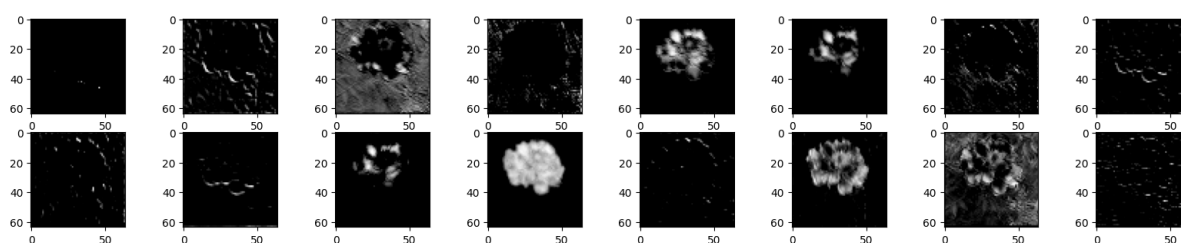


شکل ۷: تصاویر انتخاب‌شده

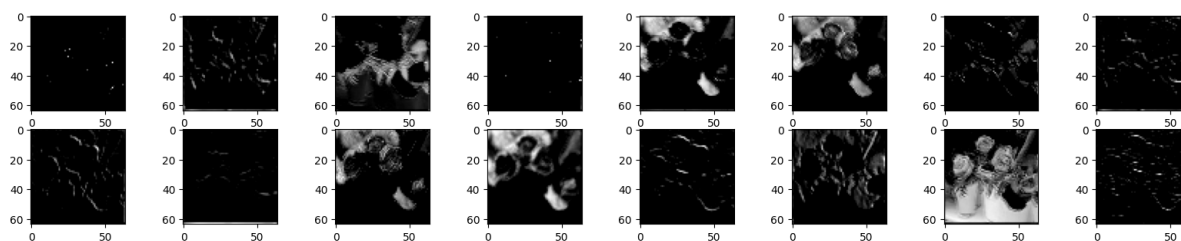
در ادامه نیز برای هر سه تصویر تعداد ۱۶ فیلتر که برای لایه اول است آورده شده است. همانطور که مشاهده می‌شود، در اکثر فیلترها لبه‌ها و مرزها بوضوح قابل رویت هستند (حتی تصویر اول که انگور است).



شکل ۸: فیلترهای تصویر انگور



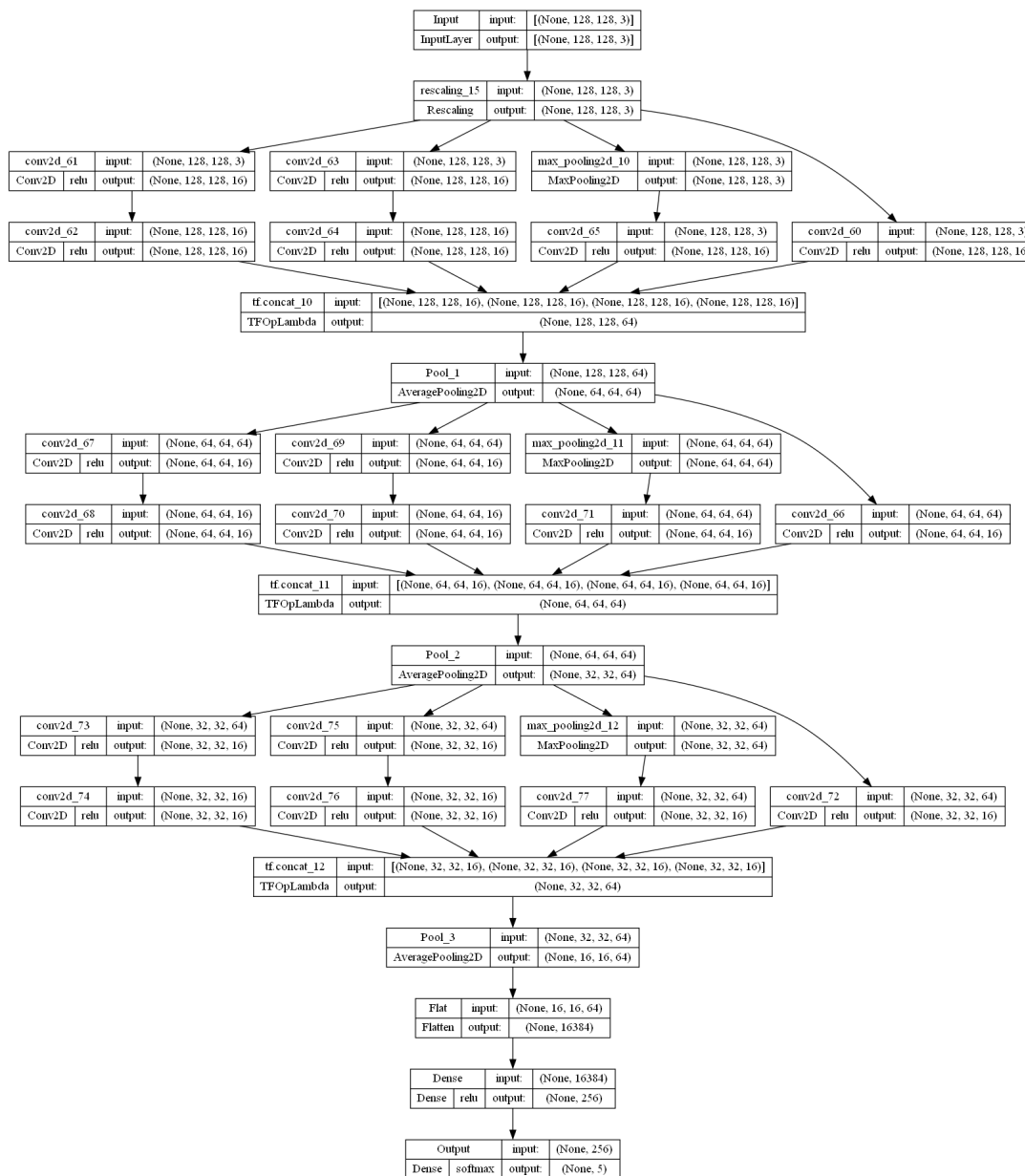
شکل ۹: فیلترهای تصویر گل نارنجی



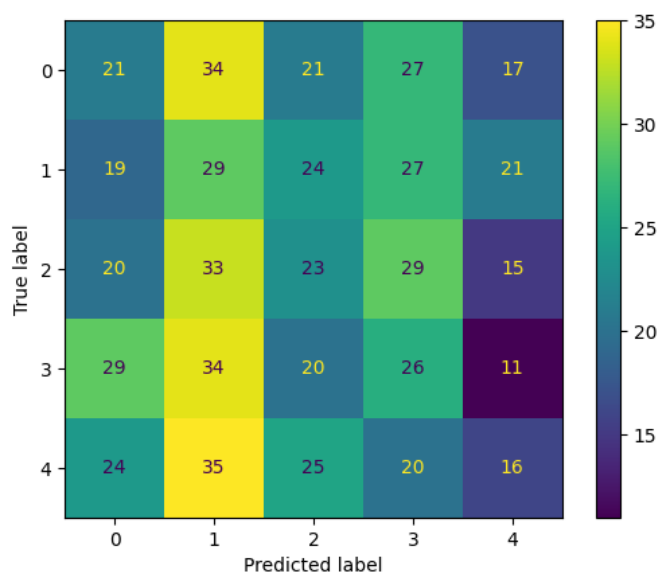
شکل ۱۰: فیلترهای تصویر گل صورتی

## ۶ پیاده‌سازی - سوال چهارم

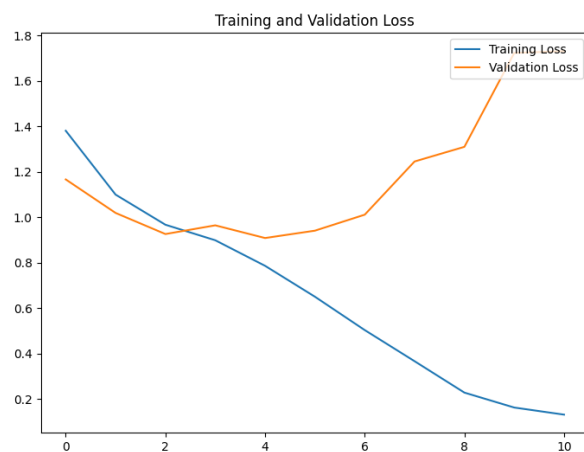
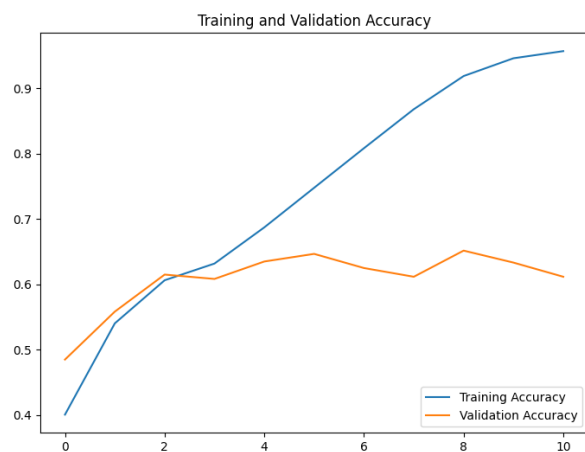
در این بخش نیاز است تا تعداد گام‌ها برابر با ۱ باشد، تا معماری کار کند. به همین جهت معماری انتخاب شده برای این بخش شامل به ترتیب ۱۶، ۶۴ و ۱۲۸ تعداد فیلتر با avg pooling و همچنین ۲۵۶ نورون لایه پنهان خواهد بود. ابتدا در ادامه معماری این قسمت مشاهده می‌شود. سپس ماتریس درهم‌ریختگی و در نهایت نمودارهای خطا و صحت نمایش داده می‌شود.



شکل ۱۱: گراف مصور مدل



شکل ۱۲: ماتریس درهم‌ریختگی



شکل ۱۳: نمودارهای خطا و صحت

اما نتایج گرفته شده جالب است. بر خلاف انتظار قبلی، در نتایج بهبودی خاصی در صحت مجموعه اعتبار سنجی وجود نداشت. بهترین صحت مجموعه اعتبار سنجی حدود ۶۵ درصد بدست آمد (در مقایسه با ۶۸ درصد در حالت اول (فایل ضمیمه)). از طرفی زمان آموزش بسیار طولانی تر شد (حدود ۱ دقیقه و ۴۵ ثانیه)، درحالیکه در حالت اولیه شبکه در حدود ۱۷ ثانیه آموزش دید. این امر به دلیل افزایش چشمگیر پارامترهای آموزشی است. ولی نکته دیگر این شبکه زود به اتمام رسیدن آموزش آن توسط کالبد در مقایسه با حالت اولیه بود. با شرایط یکسان در ای پاک ۱۰ متوقف شد (درمقایسه با ۱۲ و ۱۶ ای پاک در حالت اولیه).

## ۷ پیاده سازی - سوال پنجم

به دلیل آنکه مقادیر فیلترها با یک مجموعه آموزشی، آموزش دیده اند. به طور کلی وظایف فیلترها استخراج ویژگی ها (مانند لبه ها و مرزها) است. یک فیلتر مشخص برای هر عکس یک عملکرد مخصوصی دارد. همانطور که در سوال ۳ مشاهده شد. لذا اساسا در ابتدا نیازی به آموزش نیست. از طرفی شباهت هایی نیز بین مجموعه های داده نیز وجود دارد (وجود گل ها در مجموعه داده اول). اما برای عملکرد بهتر در گام های بعدی آموزش قسمت استخراج ویژگی ادامه می یابد.

## ۸ پیاده سازی - سوال ششم

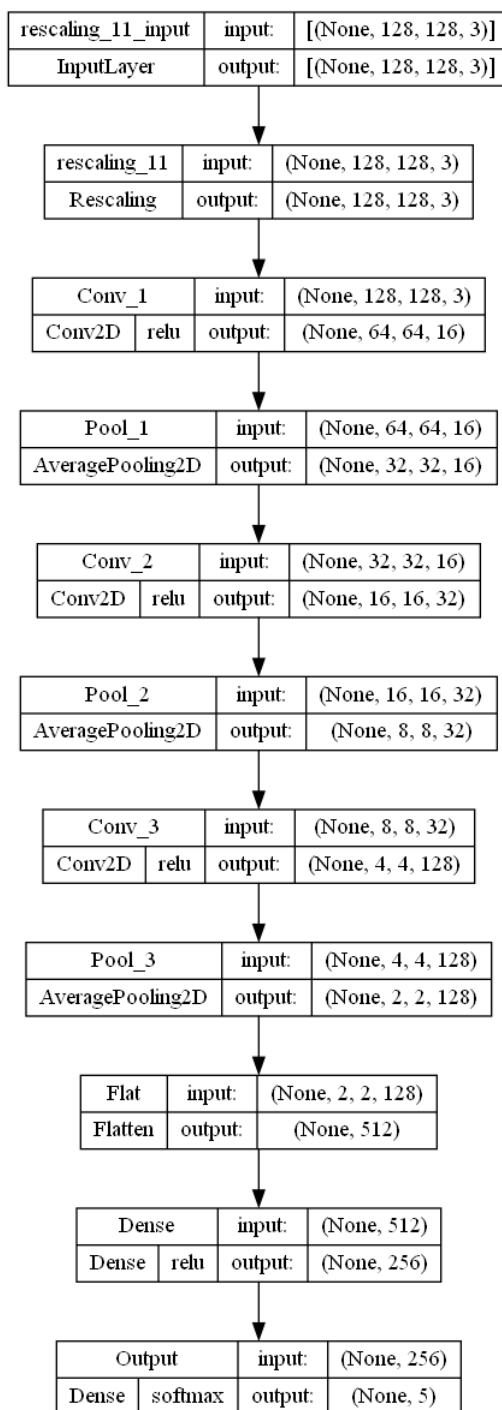
هرچه پیچیدگی دسته بند بیشتر باشد نیاز است تا در گام اول یادگیری، ای پاک های بیشتری در نظر گرفته شود. زیرا با افزایش پیچیدگی پارامترها نیز افزایش یافته و زمان بیشتری نیاز است تا آموزش ببینند. اما طراحی معماری دسته بند در این مثال خاص با روش زیر طراحی شد.

با همان روش ذکر شده از تعریف پروژه، ابتدا قسمت استخراج ویژگی را برداشته (غیر قابل آموزش) و سپس مقادیر مختلفی از تعداد نورون دسته بند با مقدار ای پاک ۱۵ آموزش می دهیم. سپس بهترین مقدار صحت مجموعه اعتبار سنجی را در نظر گرفته و مقدار متناظر نورون آن را انتخاب می کنیم. در نهایت همه پارامترهای استخراج ویژگی را قابل آموزش دهی کرده و آموزش را از سر می گیریم. جدول زیر مقدار صحت مجموعه آموزش و اعتبار سنجی به ازای تعداد نورون را نشان می دهد.

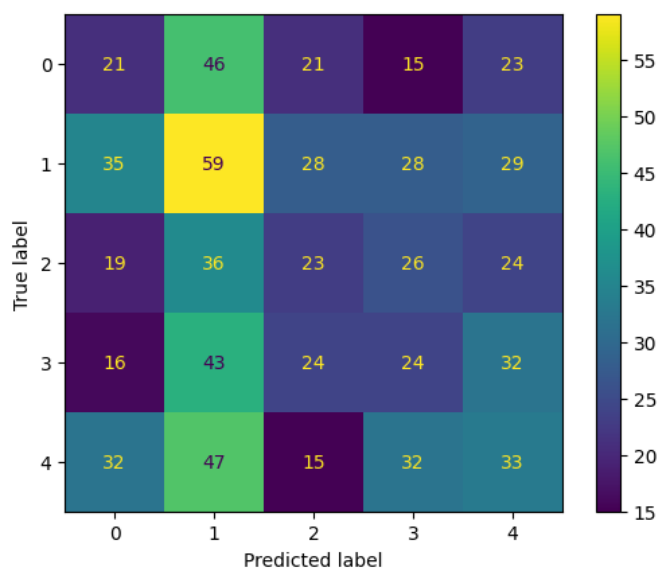
صحت تست	صحت اعتبار سنجی	صحت آموزش	نورون
43.28	46.24	43.28	۱۶
45.97	43.23	44.99	۳۲
47.58	48.7	49.71	۶۴
48.92	47.61	50.33	۱۲۸
50.27	49.25	52.12	۲۵۶

همانطور که مشاهده می‌شود بهترین تعداد نورون ۲۵۶ تا است. در ادامه آموزش را به ازای ۲۵۶ نورون تکمیل کرده و خروجی مشاهده می‌شود.

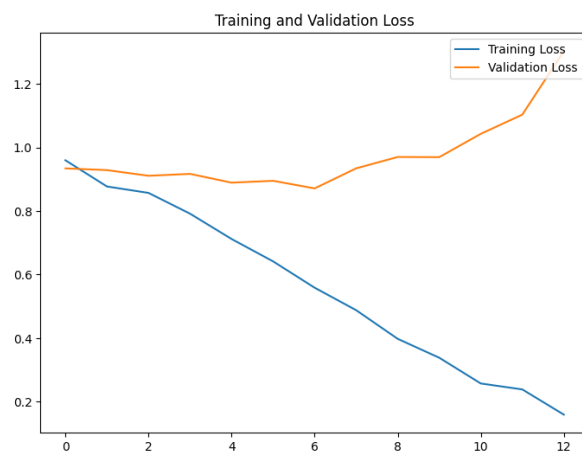
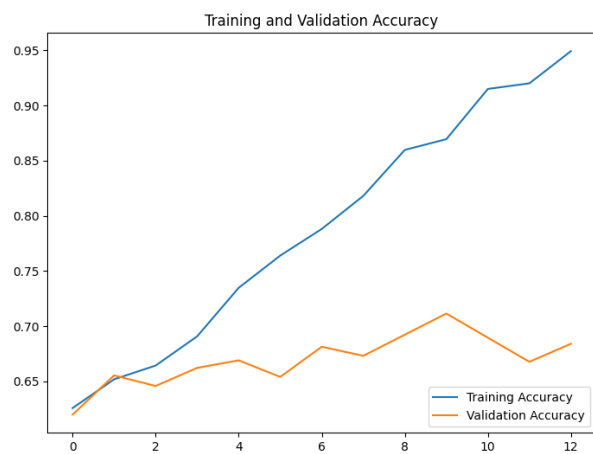
برای مدل نهایی گراف مصور مدل، ماتریس درهم‌ریختگی، نمودارهای خطا و صحت به ترتیب آورده شده‌اند.



شکل ۱۴: گراف مصور مدل



شکل ۱۵: ماتریس درهم‌ریختگی



شکل ۱۶: نمودارهای خطا و صحت

همانطور که مشاهده می‌شود، نتایج نیز در بهترین حالت در ای‌پاک ۹ با ۷۰ درصد صحت اعتبارسنجی رخ داد، که بعد از آن مدل رفته رفته دچار بیش‌برازش شد.