

در این پروژه، مدل نهایی یک شبکه CNN پنج لایه ای است که برای طبقه بندی تصاویر گربه و سگ طراحی و آموزش داده شده است. این مدل شامل پنج لایه کانولوشنی با فیلترهای افزایشی (از ۱۶ تا ۲۵۶ کانال)، به همراه لایه های Batch normalization و dropout با نرخ 0.3 و دو لایه تمام متصل در انتها می باشد. معماری به صورت end-to-end روی تصاویر خام (پس از اعمال تبدیلات افزایشی) آموزش دیده و بهترین عملکرد را در مقایسه با دو مدل دیگر نشان داده است.

برای مقایسه، دو معماری دیگر نیز پیاده سازی شدند:

- نسخه ای از CNN که از خروجی الگوریتم تشخیص لبه Canny به عنوان ورودی استفاده می کرد (به منظور تمرکز بر ویژگی های لبه ای و ساختاری).
- یک MLP ساده به عنوان مدل پایه.

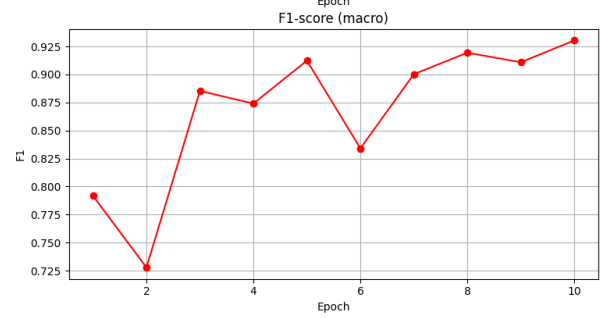
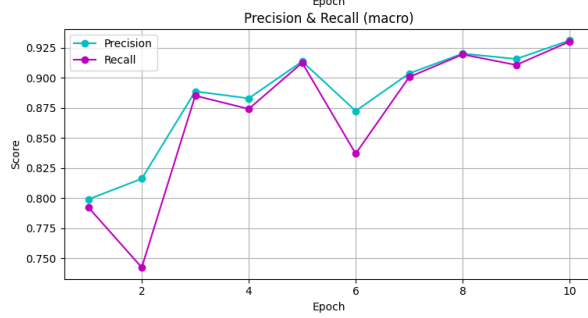
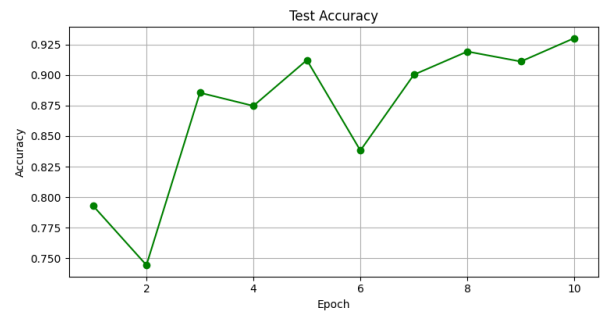
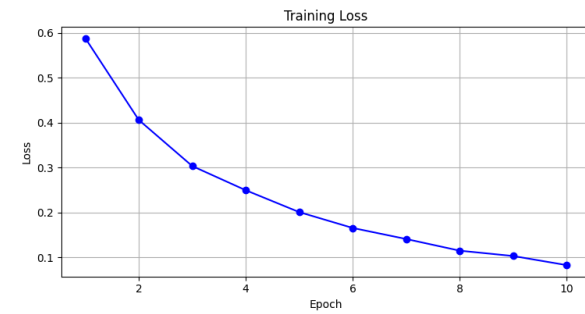
هر دو مدل جایگزین عملکرد ضعیفتری نسبت به CNN اصلی داشتند؛ MLP به دلیل عدم توانایی در استخراج ویژگی های فضایی، و نسخه Canny به دلیل از دست رفتن اطلاعات رنگی و بافتی، نتایج قابل قبولی ارائه نکردند.

مدل نهایی با استفاده از بهینه ساز ADAM و نرخ یادگیری 0.001 تابع لاس Cross-Entropy و مجموعه داده ای متعادل شامل تصاویر گربه، سگ و پس زمینه آموزش داده شد. نمودارهای کاهش زیان و افزایش دقت در طول Epoch 10، همراه با جدول دقت، دقت کلاس به کلاس (precision/recall/F1) و نرخ خطا برای مدل در گزارش ارائه شده است. تحلیل خطا نشان می دهد که بیشتر اشتباهات در تصاویر سگ های با جثه کوچک و موی گربه مانند رخ داده..

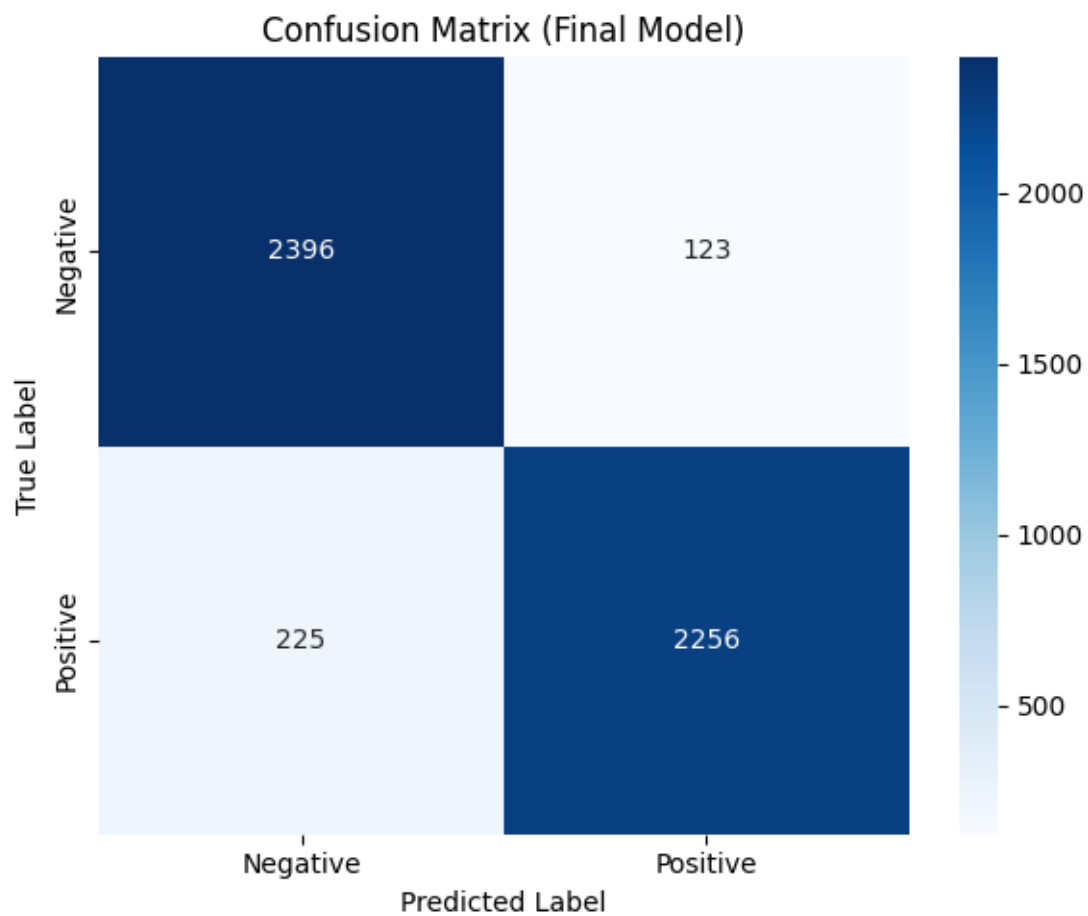
در نهایت، نتایج نشان می دهد که CNN پنج لایه ای بدون پیش پردازش Canny، با دقت حدود ۹۲٪ روی مجموعه آزمون، بهترین تعادل بین دقت، سرعت و قابلیت تعمیم را ارائه می دهد و برای کاربردهای عملی تشخیص حیوان در تصاویر مناسب است. پیشنهاد می شود در کارهای آتی از Transfer Learning با مدل های پیش آموزش دیده مانند ResNet استفاده شود تا دقت به بیش از ۹۵٪-۹۷٪ برسد. همچنین افزایش لایه های CNN می تواند مفید باشد.

در ادامه جداول و نمودار های آموزش و مدل را خواهیم دید:

Image Size	128*128
Batch Size	32
Epochs	10
Learning Rate	0.001
Optimizer	Adam
Loss Function	Cross Entropy Loss
Convolutional Layers	5
Input Channels	1(Greyscale)
Data Augmentation	RandAugment and many others
Dropout Rate	0.3
Regularization	L2
Weight Decay	0.0001



نمودار های Accuracy، Loss، F1score و Precision و Recall به صورت Macro.

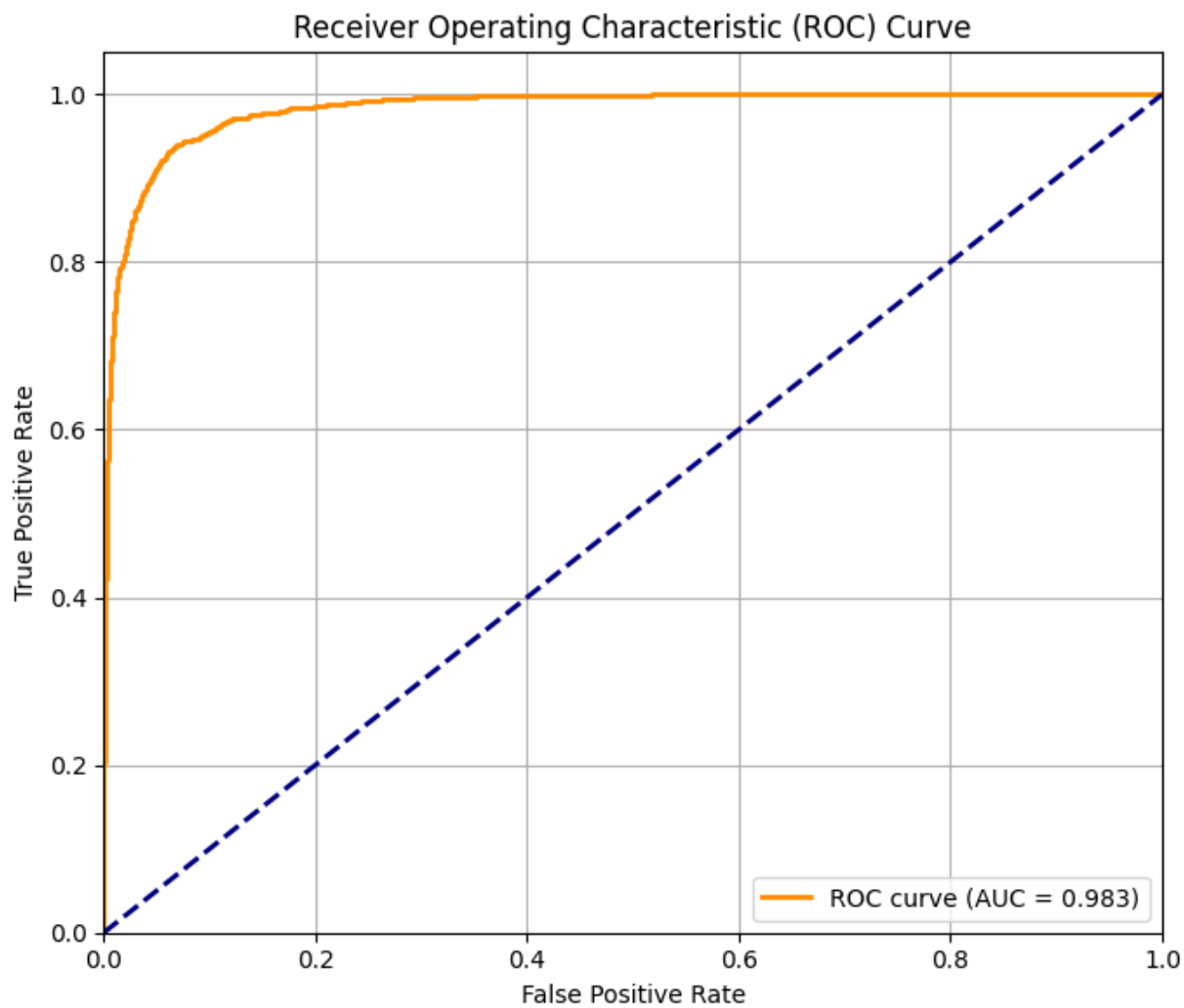


Confusion Matrix

Positive = Dog

Negative = Cat

مدل در تشخیص تصاویر سگ تقریباً 2 برابر ضعیف تر از تشخیص گربه ها عمل می کند.



## Cat vs Dog CNN



Browse Image

Prediction: DOG (0.61)