

در این پروژه، مدل نهایی یک شبکه CNN پنج لایه‌ای است که برای طبقه‌بندی 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 لایه‌های Batch normalization و dropout با نرخ ۰.۳ و دو لایه تمام‌متصل در انتهای می‌باشد. 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 پیاده‌سازی شدند:

- نسخه‌ای از CNN که از خروجی الگوریتم تشخیص لبه Canny به عنوان ورودی استفاده می‌کرد (به منظور تمرکز بر ویژگی‌های لبه‌ای و ساختاری).
- یک MLP ساده به عنوان مدل پایه.

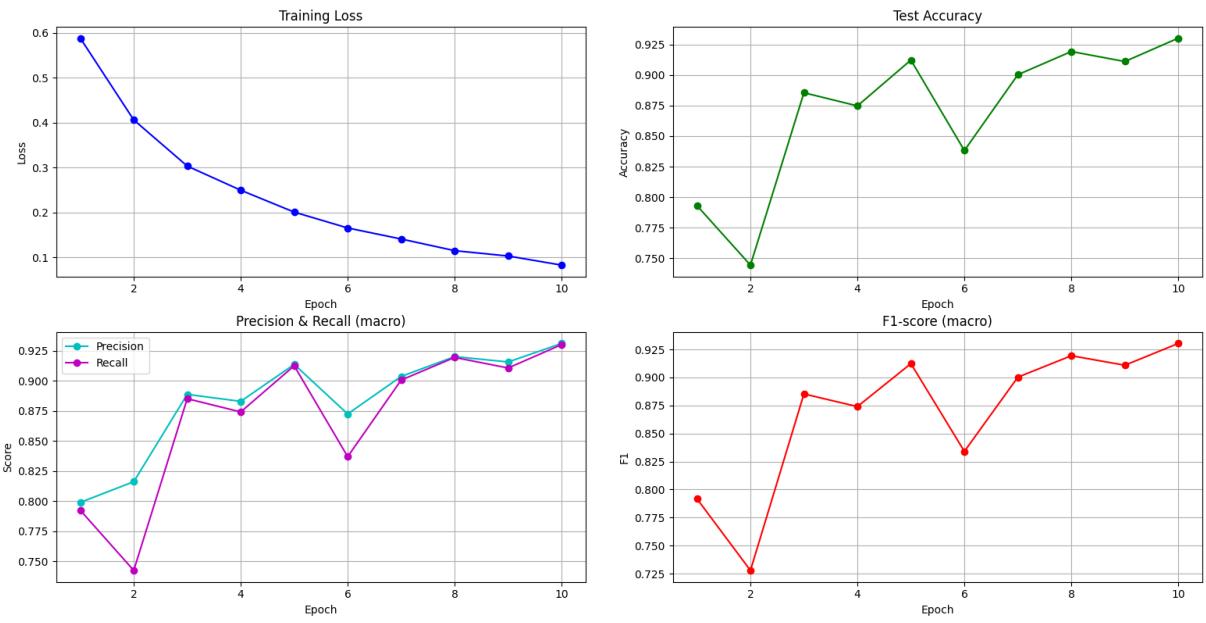
هر دو مدل جایگزین عملکرد ضعیفتری نسبت به CNN اصلی داشتند؛ MLP به دلیل عدم توانایی در استخراج ویژگی‌های فضایی، و نسخه Canny به دلیل از دست رفتن اطلاعات رنگی و بافتی، نتایج قابل قبولی ارائه نکردند.

مدل نهایی با استفاده از بهینه‌ساز ADAM و نرخ یادگیری ۰.۰۰۱ Cross-Entropy تابع لاس و مجموعه داده‌ای متعادل شامل تصاویر گربه، سگ و پس زمینه آموزش داده شد. نمودارهای کاهش زیان و افزایش دقت در طول ۱۰ Epoch، همراه با جدول دقت، دقت کلاس به کلاس (precision/recall/F1) و نرخ خطأ برای مدل در گزارش ارائه شده است. تحلیل خطأ نشان می‌دهد که بیشتر اشتباهات در تصاویر سگ‌های با جثه کوچک و موی گربه مانند رخ داده.

در نهایت، نتایج نشان می‌دهد که CNN پنج لایه‌ای بدون پیش‌پردازش Canny، با دقت حدود ۹۲٪ روی مجموعه آزمون، بهترین تعادل بین دقت، سرعت و قابلیت تعمیم را ارائه می‌دهد و برای کاربردهای عملی تشخیص حیوان در تصاویر مناسب است. پیشنهاد می‌شود در کارهای آنی از Transfer Learning با مدل‌های پیش‌آموزش دیده مانند ResNet استفاده شود تا دقت به بیش از ۹۷٪–۹۵٪ برسد. همچنین افزایش لایه‌های CNN می‌تواند مفید باشد.

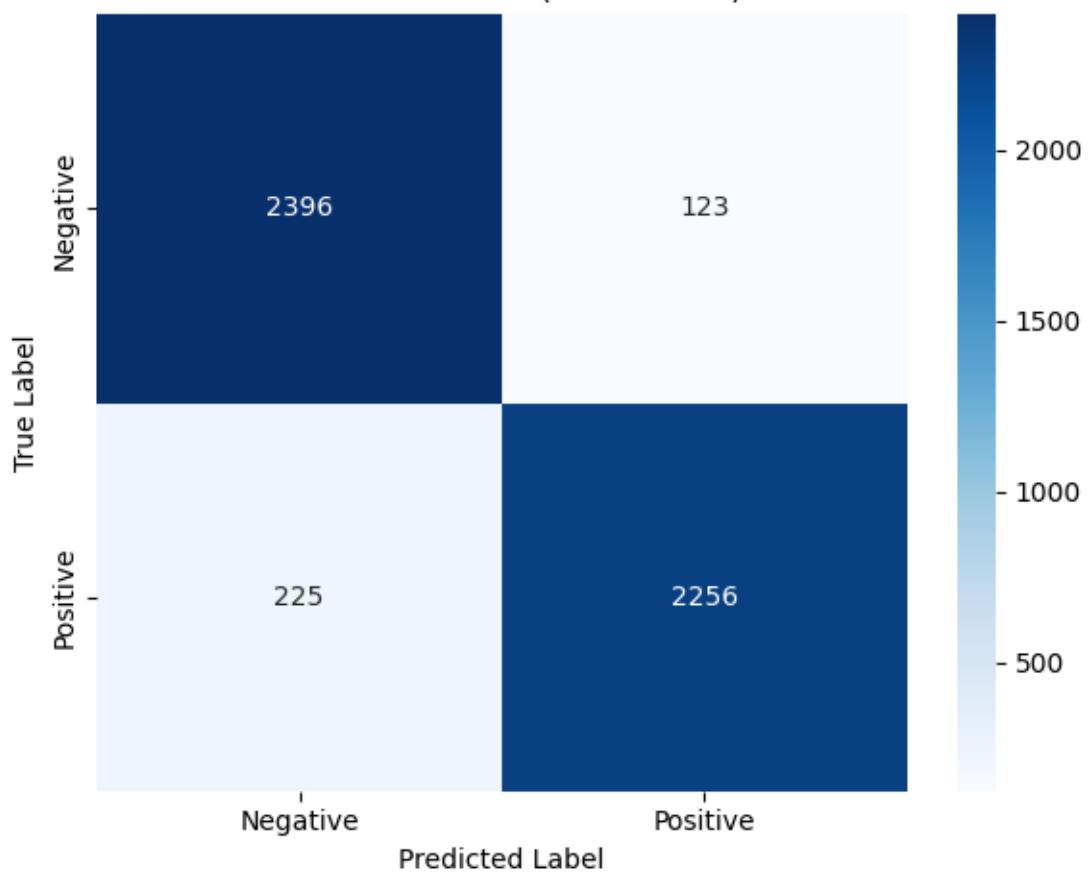
در ادامه جداول و نمودارهای آموزش و مدل را خواهیم دید:

Image Size	128*128
Batch Size	32
Epochs	10
Learning Rate	0.001
Optimizer	Adam
Loss Function	Cross Entropy Lossss
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 Loss، Accuracy، Precision و Recall به صورت Macro

Confusion Matrix (Final Model)

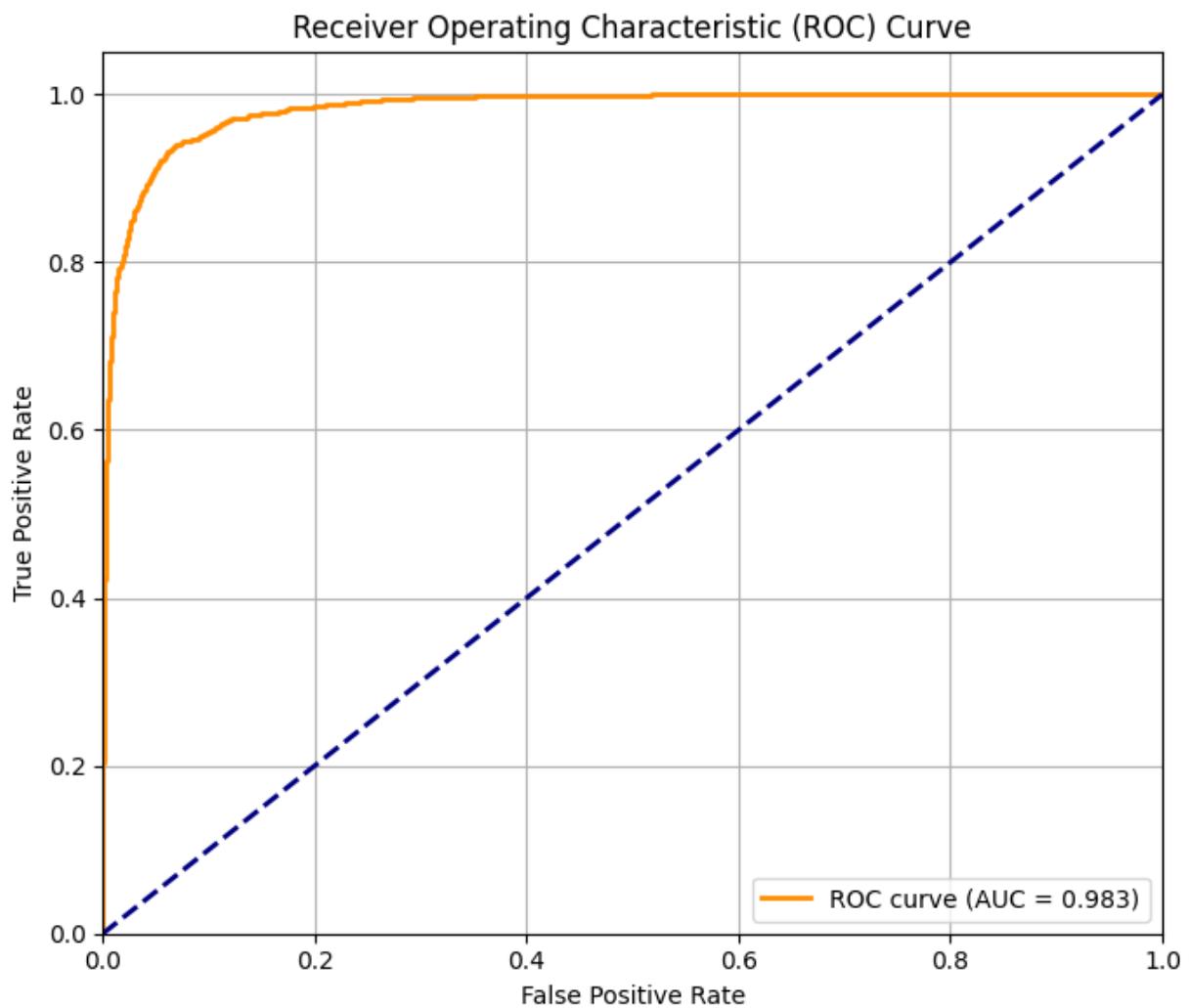


Confusion Matrix

Positive = Dog

Negative = Cat

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





Cat vs Dog Classifier (PyTorch)



Cat vs Dog CNN



Browse Image

Prediction: DOG (0.61)