

من تو این پروژه با دیتاست CIFAR-10 کار کردم که شامل ۱۰ کلاس مختلف از تصاویر رنگی با اندازه‌ی  $32 \times 32$  هست. برای آموزش، از روش‌های پیش‌پردازش استاندارد مثل نرمال‌سازی با mean/std مخصوص CIFAR-10 استفاده کردم و برای بهتر شدن تعمیم‌پذیری مدل هم روی داده‌های آموزشی RandomCrop و padding با RandomHorizontalFlip گذاشتم. هدفم این بود که هم یک مدل طبقه‌بندی قابل قبول بسازم و هم بتونم رفتار لایه‌های مختلف شبکه رو به شکل قابل مشاهده تحلیل کنم.

برای مدل هم از یک CNN دست‌ساز سبک VGG-style استفاده کردم که شامل چند بلاک کانولوشن (جمعاً ۸ لایه‌ی) به همراه MaxPool و ReLU و BatchNorm با momentum و Conv weight decay کم کردم. نتیجه‌ی آموزش خیلی قابل قبول شد و دقت اعتبارسنجی به حدود ۹۲٪ رسید.

تو این تمرین Feature Map‌های چند تا لایه‌ی مختلف CNN رو روی CIFAR-10 برسی کردم. چیزی که خیلی جالب بود اینه که شبکه از لایه‌های اول با چیزی خیلی ساده شروع می‌کنه و هرچی جلوتر می‌ریم، ویژگی‌هایی که یاد می‌گیره پیچیده‌تر و نزدیکتر به مفهوم «کلاس» می‌شن. یعنی لایه‌های اول بیشتر مثل یک فیلتر ساده عمل می‌کنن، ولی لایه‌های عمیق‌تر انگار روی بخش‌های مهم تصویر تمرکز می‌کنن.

۱) لایه‌های اولیه چه چیزی را یاد می‌گیرند؟

تو لایه‌های اولیه بیشتر خروجی‌ها شبیه لبه‌ها و خطوط هستن. مثلاً خطوط افقی/عمودی/مورب و جاها‌یی که تغییر روشنایی یا تغییر رنگ شدید داریم واضح‌تر دیده می‌شن. به نظرم این لایه‌ها بیشتر دارند «اطلاعات خام» تصویر رو به شکل قابل استفاده‌تر تبدیل می‌کنن.

۲) ویژگی‌های لایه‌های میانی چه فرقی با لایه‌ی اول دارند؟

وقتی می‌ریم لایه‌های میانی، دیگه فقط لبه و خط نیست. اینجا شبکه شروع می‌کنه چند تا ویژگی ساده رو با هم ترکیب کردن و چیزهایی مثل بافت‌ها، الگوهای تکراری، گوشه‌ها و قسمت‌های کوچیک از شکل‌ها رو بهتر نشون می‌ده. یعنی لایه‌های میانی مثل اینکه از روی لبه‌ها، یه جور «الگوهای معنی‌دارتر» بسازن.

(3) در لایه‌های عمیق‌تر چه اتفاقی می‌افتد؟

تو لایه‌های عمیق‌تر Feature Map ها معمولاً انتخابی‌تر می‌شون. یعنی خیلی از کانال‌ها ممکن‌هست کم فعال باشند، ولی چند تا کانال روی قسمت‌های خاصی از تصویر خیلی واضح فعال می‌شون. برداشت من اینه که اینجا شبکه بیشتر دنبال چیزاییه که برای تشخیص کلاس واقعاً مهمه و سعی می‌کنه به جزئیات اضافه مثل پس‌زمینه کمتر توجه کنه.

(4) اگر تصویر ورودی عوض شود، آیا الگوی مشابهی در لایه‌ها دیده می‌شود؟

به نظرم تو لایه‌های اول تا حد زیادی الگوها شبیه هم می‌مونن، چون لبه و کنتراست تو همه‌ی تصاویر وجود داره. ولی هرچی لایه‌ها عمیق‌تر می‌شون، خروجی‌ها بیشتر وابسته به خود تصویر و کلاسش می‌شون و تفاوت‌ها واضح‌تر دیده می‌شه. یعنی تغییر ورودی تو لایه‌های عمیق اثر بیشتری می‌ذاره.