**کتاب پیاده سازی یادگیری ماشین در بینایی ماشین**

**بخش سه: دید تصویر Image Vision**

در این بخش مفاهیم و تکنیک های دسته بندی تصاویر، کانولوشن، ترنسفر لرنینگ و فاین تیونینگ طرح می شود.

* منظور از ترنسفر لرنینگ این است که مدلی که از قبل روی تعداد زیادی تصویر (بالای 5000 تصویر بازای هر لیبل) آموزش دیده است را استفاده کنیم تا روی داده های ما دسته بندی انجام دهد. در این روش ضرایب مربوط به لایه های شبکه عصبی بصورت کامل دانلود می شود و فقط لایه آخر که نیاز به انجام دسته بندی در کاربرد جدید و روی دیتاست جدید می باشد تغییر می کند و ضرایب آن تغییر می کند.
* منظور از فاین تیونینگ این است که ضرایب مدل پیش آموزش دیده Pretrained بعنوان مقادیر اولیه در نظر گرفته می شوند و لایه های شبکه عصبی سعی در بهبود آن می کنند. این در حالی است که در تعریف شبکه های عصبی مقادیر بصورت رندوم در نظر گرفته می شوند و سپس بهبود می یابند.
* انتخاب استراتژی در استفاده از روش های مختلف:
  + اگر دیتاست کوچک (کمتر از 1000 داده بازای هر لیبل) داشته باشیم، بهتر است از ترنسفر لرنینگ استفاده کنیم.
  + اگر دیتاست متوسط داشته باشیم (بین 1000 تا 5000 داده بازای هر لیبل) بهتر است از فاین تیونینگ استفاده کنیم.
  + اگر دیتاست بزرگ داشته باشیم (بیش از 5000 داده بازای هر لیبل) بهتر است مدل را از ابتدا scratch آموزش دهیم.
* چه مدلی را انتخاب کنیم:
  + اگر بخواهیم تعریف لایه ها و ساختار آن در اختیار خودمان باشد، بهتر است از یک مدل ساده مثل SqueezeNet که کارآمدی خوبی دارد استفاده شود.
  + اگر بخواهیم از یک مدل سبک و سریع که حافظه کمی استفاده کند و تاخیر محاسبات نداشته باشد، بهتر است از مدل MobileNetV2 استفاده شود.
  + اگر محدودیتی در زمان و هزینه محاسبات و غیره نداشته باشیم و فقط دقت مدل اهمیت زیادی داشته باشد، بهتر است از مدل EfficientNet استفاده کنیم.
  + اگر صحت و دقت مدل بسیار اهمیت دارد و محافظه کاری در دستور کار باشد، بهتر است از مدل ResNet50 یا بالاتر استفاده شود.