

به نام خدا

در این بخش قسمت ت را توضیح میدهم و در pdf گزارش پروژه را میدهم.

## بخش ت

در پروژه ای که پیاده سازی کردم لایه تمام متصل یکی مانده به آخر از شبکه RESNET را برای استخراج ویژگی انتخاب کردم.

کدش به این شکل است:

```
resnet50 = models.resnet50(pretrained=True)
```

```
feature_extractor =
```

```
nn.Sequential(*list(resnet50.children())[:-1])
```

در این قسمت ما همه لایه های resnet را فراخوانی کردیم و به یک لیست تبدیل کردیم. لایه آخر را ازش جدا کردیم و یک مدل جدید ساختیم.

بر اساس ساختار شبکه resnet ۲ لایه آخر به این شکل هستند:

```
(avgpool): AdaptiveAvgPool2d(output_size=(1, 1))  
(fc): Linear(in_features=2048, out_features=1000, bias=True)
```

بنابراین لایه انتخابی برای استخراج ویژگی لایه avgpool است.

چرا این لایه؟

این لایه چون به خروجی شبکه عصبی نزدیک تر است ویژگی های معنا دار تری و سطح بالاتری استخراج کرده است نسبت به لایه های ابتدایی که ویژگی های ابتدایی را ذخیره کرده است.

این لایه خیلی عمیق است در شبکه رزنت و اطلاعات سطح بالایی استخراج کرده است.

همچنین این لایه یک **average pooling** در سراسر **spatial** ما در **feature map** اجرا کرده است که باعث میشود کل ورودی ما خلاصه سازی شود.

همچنین این لایه ابعاد آن ثابت شده است و دیگر تاثیری از ورودی نمیپذیرد. این نکته خیلی مهم است زیرا لایه های تمام متصل نیازمند سایز ثابت هستند.

از هر فریم چه مقدار ویژگی استخراج کردیم؟

تعداد ویژگی استخراج شده برای هر فریم برابر با ۲۰۴۸ است.

بعد از انجام **avgpool** بعد بردار ویژگی ما به ۲۰۴۸ ثابت میشود.

چه راهی برای کاهش تعداد ویژگی پیشنهاد میتوانیم بکنیم؟

روش های مختلفی وجود دارد برای کاهش تعداد ویژگی:

اولین روش **PCA** است که بعد کاهش خطی بعد بسیار مناسب است. که با استفاده از پیدا کردن مولفه های اصلی این اتفاق میفتد. ما میتوانیم تعداد کوچکی از مولفه های را انتخاب کنیم که بیشترین میزان اطلاعات را برمیگردانند و باعث کاهش بعد میشوند.

دومین روش مطرح، **autoencoder** ها هستند که آموزش دیده اند برای بازسازی ورودی که بهشان داده میشود. آن ها یک گلوگاه (**bottleneck**) دارند که بسیار کوچکتر از ورودی است. این فعالیت این لایه باعث کاهش بعد میشود.

سومین روش اضافه کردن یک لایه تمام متصل با تعداد نورون های کمتر هست در لایه قبل از طبقه بندی ما که باعث میشود خروجی ما مجبور شود که فشرده شود.