

گزارشی مختصر از پروژه تشخیص لبخند درس یادگیری عمیق

آرمان خلیلی - ۴۰.۳۶۲۳۰.۱۶

## مراحل انجام پروژه:

(۱) دانلود و آماده سازی دیتاست

دانلود دیتاست از Kaggle

(۲) انجام Face detection

با استفاده از مدل تشخیص چهره YuNet استخراج صورت ها از تصاویر انجام گرفت. این مدل در عین بهینگی و داشتن تنها ۷۵۸۵۶ پارامتر دقت بسیار خوبی در یافتن چهره های تصویر فراهم میکند. این مدل را پس از امتحان مدل های متعدد و خواندن چند مطلب انتخاب کردم. یکی از این مطالب در زیر معرفی شده است:

[What's the Best Face Detector-by Amos Stailey-Young - Python's Gurus - Medium](#)

(۳) پیش پردازش تصاویر

در این گام پیش پردازش تصاویر چهره ها طی چند مرحله انجام گرفت. این مراحل عبارتند از تبدیل تصاویر BRG به RGB، تغییر سایز تصاویر به  $۱۶۰ \times ۱۶۰$ ، نرمالایز کردن و نگاشت مقادیر ماتریس به بازه ۰ تا ۱.

(۴) Fine-tuning a pre-trained Image-classification Model

در این مرحله یک مدل کلاسبندی تصاویر که قبلا روی یک دیتاست بزرگ آموزش دیده را انتخاب میکنیم و برای تسک خودمان باز تنظیم کرده و چند لایه جدید به انتهای آن اضافه کرده و آموزش میدهیم. مدل انتخاب شده VGG19 هست که این مدل نیز پس از امتحان چند مدل مختلف مانند InceptionV3 و Resnet50 و داشتن عملکرد بهتر و لایه های کمتر انتخاب کردم. خواندن مقاله زیر نیز در انتخاب این مدل به من کمک کرد:

[Exploring the Potential of Pre-trained DCNN Models for Facial Emotion Detection- A Comparative Analysis](#)

## ۵) اجرای مدل روی تصویر Webcam بصورت زنده

در این مرحله این پس از اتمام آموزش مدل، آنرا دانلود کرده و یک کد آماده کردم که تصویر وبکم را بصورت فریم به فریم ابتدا تشخیص چهره بدهد و سپس تصویر پیش پردازش شده را به مدل تشخیص لبخند بدهد و در لحظه روی تصویر وضعیت چهره پیشبینی شده را روی تصویر بنویسد.

### معماری و دقت مدل تشخیص لبخند:

مدل VGG19 که در سال ۲۰۱۴ معرفی شد و از ۱۹ لایه conv و ۵ لایه MaxPooling تشکیل شده و دارای ۲۰ میلیون پارامتر است. ما در این تسک تا لایه ۱۷ را فیریز کردیم. همچنین دو لایه Dense در انتهای آن اضافه کردیم. در جدول زیر معماری مدل نهایی را مشاهده میکنید:

Model: "functional"

Layer (type) #	Output Shape	Param
input_layer_1 (InputLayer)	(None, 160, 160, 3)	0
vgg19 (Functional)	(None, 5, 5, 512)	20,024,384
global_average_pooling2d (GlobalAveragePooling2D)	(None, 512)	0
batch_normalization (BatchNormalization)	(None, 512)	2,048
dense (Dense)	(None, 512)	262,656
dropout (Dropout)	(None, 512)	0

batch_normalization_1 (BatchNormalization)	(None, 512)	2,048
dense_1 (Dense)	(None, 256)	131,328
dropout_1 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 1)	257

Total params: 20,422,721 (77.91 MB)  
 Trainable params: 9,835,521 (37.52 MB)  
 Non-trainable params: 10,587,200 (40.39 MB)

دقت مدل:

Validation Loss: 0.17929725348949432  
 Validation Accuracy: 0.9466328024864197

نمودارهای یادگیری:

