

## به نام خدا

### یادگیری ماشین

#### شبکه های عصبی مصنوعی استخراج ویژگی های تصویر

<http://www-2.cs.cmu.edu/afs/cs.cmu.edu/user/mitchell/ftp/faces.html>

#### مقدمه

این تمرین به شما امکان می دهد تا از روش یادگیری شبکه های عصبی مصنوعی برای حل یک مسأله واقعی استفاده کنید. انجام این تمرین مستلزم برنامه نویسی زیادی نیست و بخش اصلی برنامه نوشته شده است. وظیفه شما فهم منطق کلی برنامه و توانایی پاسخگویی به سوالات مطرح شده درباره مراحل کار است. تسلط بر تمام کد لازم نیست.

#### تصویرها

شاخه faces\_4 حاوی فایل های تصاویر است. در این شاخه ۲۰ زیر شاخه وجود دارد که هر یک مربوط به یکی از افراد است. این فایلها پسوند pgm دارند و توسط ACDSee قابل مشاهده اند. نام هر فایل نشان دهنده ویژگی های تصویر است و به شکل زیر است:

pgm . <scale><eyes><expression><pose><userid>

- <userid> شناسه یک از ۲۰ نفر است، مانند an2i ، mitchel ، sz24 و...
- <pose> جهت صورت است و یکی از ۴ مقدار left ، right ، straight و up را خواهد داشت.
- <expression> حالت چهره است و یکی از ۴ مقدار happy ، sad ، neutral و angry را خواهد داشت.
- <eyes> داشتن یا نداشتن عینک را نشان می دهد : open یا sunglasses
- <scale> مقیاس تصویر است. در این تمرین مقدار آن ثابت و مساوی ۴ است.

#### مراحل کار

۱. برنامه به زبان سی و برای لینوکس نوشته شده است. این کد در حال حاضر تشخیص می دهد که آیا تصویر مربوط به glickman است یا خیر. برنامه را طوری تغییر دهید که داشتن یا نداشتن عینک را تشخیص دهد. یعنی یک شبکه عصبی بسازید که تصویر را به عنوان ورودی بگیرد و تشخیص دهد که آیا چهره موجود در تصویر عینک دارد یا خیر.
۲. شبکه عصبی را به کمک دستور زیر آموزش دهید :  
./facetrain -n shades.net -t ../list/straightrnd\_train.list
- پارامترهای لازم برای facetrain در بخشهای بعدی توضیح داده می شود. این دستور شبکه عصبی را به ازای تصاویر موجود در لیست straightrnd\_train.list آموزش می دهد. شبکه عصبی حاصل، در فایل shades.net ذخیره می شود.
۳. حال می خواهیم ببینیم شبکه آموزش یافته تا چه حد درست کار می کند. برای این کار شبکه را روی تصاویر موجود در دو فایل straightrnd\_test1.list و straightrnd\_test2.list آزمایش می کنیم. دقت کنید که تصاویر موجود در این دو لیست در لیست straightrnd\_train.list وجود ندارند(چرا نباید باشند؟)

./facetrain -n shades.net -T 1 ../list/straightrn\_test1.list -2 ../list/straightrn\_test2  
با اجرای این دستور تصاویری که شبکه در تشخیص عینکی بودن آنها اشتباه کرده، مشخص می شوند.

۴. خروجی شبکه عصبی را برای تصاویر مختلف مشاهده کنید، مثلاً :  
./facetrain -n shades.net -R ../faces\_4/an2i/an2i\_left\_angry\_sunglasses\_4.pgm

۵. کد را طوری تغییر دهید که تصویر را به عنوان ورودی بگیرد و تشخیص دهد کدامیک از ۲۰ نفر است و نام او را چاپ کند. برای این منظور تعداد نورونهای لایه میانی را افزایش دهید (مثلاً ۲۰ نورون)

۶. شبکه عصبی را آموزش دهید و آن را بیازمایید.

۷. کد را طوری تغییر دهید که تصویر را به عنوان ورودی بگیرد و جهت صورت را تشخیص دهد (left, right, up, down). تعداد نورونهای لایه میانی را ۶ و ثابت momentum را ۰,۲ قرار دهید. (راستی ثابت momentum یا همان اینرسی چه بود؟)

۸. شبکه عصبی را آموزش دهید و آن را بیازمایید.

۹. اگر تمام مراحل بالا را انجام دادید و فهمیدید که چه کار کردید، کار تمام است!

### فایل های موجود در پروژه

- Pgmimage.h و Pgmimage.c : توابع کار با فایل های از نوع pgm. تغییر در این فایلها لازم نیست.
- Backprop.h و backprop.h : الگوریتم پس انتشار خطا. تغییر در این فایلها لازم نیست ولی دانستن منطق پس انتشار خطا لازم است.
- Imagenet.c : تغییر در متد load\_target() این برنامه لازم است.
- Facetrain.c : این فایل از توابع فایل های دیگر استفاده می کند و برای پیاده سازی مراحل مختلف تغییر خواهد کرد.
- Hidtopgm.c : برای نمایش گرافیکی وزنه های شبکه عصبی به کار می رود ولی جزو این تمرین نیست.

### اجرای facetrain

در اینجا به بررسی پارامترهای ممکن برنامه facetrain می پردازیم. هر یک از این پارامترها توسط یکی از کلیدهای زیر مشخص می شود :

- **-n** : نام فایلی که شبکه عصبی در آن ذخیره می شود را مشخص می کند. این فایل می تواند ورودی یا خروجی باشد. در زمان یادگیری این پارامتر فایل خروجی را مشخص می کند و در زمان استفاده از شبکه عصبی فایل حاوی وزنه های شبکه عصبی را مشخص می کند.
- **-e** : تعداد epoch ها
- **-T** : شبکه عصبی در حالت آزمایش است. شبکه عصبی روی هر یک از دو لیست مشخص شده توسط پارامترهای 1- و 2- اجرا می شود و تصاویری که به اشتباه طبقه بندی شوند مشخص خواهند شد.
- **-t** : لیست تصاویر لازم برای آموزش شبکه عصبی را مشخص می کند. این کلید نمی تواند به همراه کلید T- به کار رود. (چرا؟)
- **1- و 2-** : دو فایل برای تست کارایی شبکه عصبی آموزش داده شده. فایلهایی که توسط این دو کلید مشخص می شوند و شبکه عصبی آنها را به اشتباه طبقه بندی می کند، نمایش داده خواهند شد.
- **-R** : خروجی شبکه عصبی برای یک تصویر خاص که توسط این کلید مشخص می شود، نمایش داده خواهد شد.