



## پروژه چهارم

**هدف:** آشنایی با شبکه های عصبی پیچشی.

**کد:** پیاده سازی این پروژه را به زبان پایتون انجام دهید؛ در این فعالیت مجاز به استفاده از tensorflow یا pytorch می باشید.  
**گزارش:** ملاک اصلی انجام فعالیت، گزارش آن است و ارسال کد بدون گزارش فاقد ارزش است. برای این فعالیت یک فایل گزارش در قالب pdf تهیه کنید و در آن برای هر سوال، تصاویر ورودی، تصاویر خروجی و توضیحات کامل و جامعی تهیه کنید.

**تذکر ۱:** مطابق قوانین دانشگاه هر نوع کپی برداری و اشتراک کار دانشجویان غیر مجاز بوده و شدیداً برخورد خواهد شد. استفاده از کدها و توضیحات اینترنت به منظور یادگیری بلامانع است، اما کپی کردن غیرمجاز است.

**تذکر ۲:** مجموعه های داده مورد استفاده را به جز در مواردی که صریحاً در صورت سوال ذکر شده باشد، حتماً قبل از استفاده بصورت تصادفی به سه بخش آموزش (۷۰ درصد داده ها)، آزمون (۲۰ درصد داده ها) و اعتبارسنجی (۱۰ درصد داده ها) تقسیم نمایید.

**تذکر ۳:** برای تمامی مدل های آموزش داده شده خود ماتریس درهم ریختگی، گراف مصور مدل، نمودار تغییرات خطا و صحت پیش بینی را گزارش و تحلیل کنید.

**راهنمایی:** در صورت نیاز میتوانید سوالات خود را در خصوص پروژه از تدریسارهای درس، از طریق ایمیل زیر یا در گروه تلگرامی بپرسید.

Email: ann.ceit.aut@gmail.com

**توجه:** برای آموزش شبکه های عمیق می توانید از منابع و بسترهای سخت افزاری برخط رایگان نظیر Google Colab یا Kaggle استفاده نمایید.

**تاخیر مجاز:** در طول ترم، مجموعاً مجاز به حداکثر ده روز تاخیر برای ارسال تمرینات هستید (بدون کسر نمره). این تاخیر را می توانید بر حسب نیاز بین تمرینات مختلف تقسیم کنید؛ اما مجموع تاخیرات تمام تمرینات شما نباید بیشتر از ده روز شود. پس از استفاده از این تاخیر مجاز، هر روز تاخیر باعث کسر ۱۰٪ نمره ی آن تمرین خواهد شد.

**ارسال:** فایل های کد و گزارش خود را در قالب یک فایل فشرده با فرمت StudentID\_HW04.zip تا تاریخ ۱۴۰۱ / ۰۹ / ۱۸ ارسال نمایید.

**به هر یک از سوالات زیر حداکثر در دو بند پاسخ تشریحی خود را ارائه دهید:**

- i. اتصال باقیماندگی<sup>۱</sup> در شبکه های عصبی (پیچشی) برای رفع مشکل پس انتشار خطا به لایه های ابتدایی معرفی و ارائه شده است. با ذکر دلیل و استدلال خود بیان کنید که در چه شرایطی طول اتصال باقیماندگی بهتر است بلند باشد و در چه شرایطی کوتاه؟ به چه عواملی بستگی دارد؟
- ii. در شبکه های عصبی پیچشی متراکم<sup>۲</sup>، به نظر شما دلیل استفاده از average pooling به جای max pooling در لایه ی انتقال<sup>۳</sup> چیست؟

<sup>1</sup> Residual Connection

<sup>2</sup> Densely Connected Convolution Neural Network

<sup>3</sup> Transition Layer

iii. در پروژه دوم ملاحظه کردید که طراحی یک شبکه‌ی عصبی چندلایه‌ی پرسپترون برای دسته‌بندی<sup>۴</sup> چه چالش‌هایی داشته و با تکیه بر سعی و خطا می‌توان به یک مدل مناسب دست یافت. حال در شبکه‌های عصبی پیچشی علاوه بر طراحی دسته‌بند، بایستی قسمت استخراج ویژگی<sup>۵</sup> نیز طراحی شود. به نظر شما روند طراحی توامان دسته‌بند و استخراج ویژگی بایستی چگونه انجام پذیرد تا بتوان به یک مدل عصبی پیچشی مناسب دست یافت؟

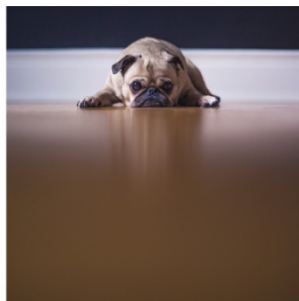
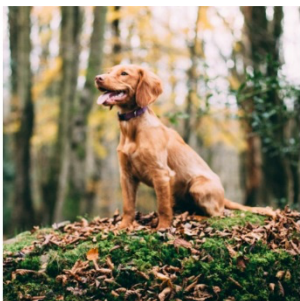
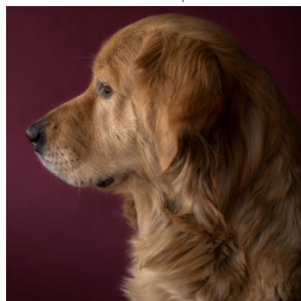
## دسته‌بندی تصاویر

مجموعه داده‌ی چالشی Linnaeus ۵ نسخه ۱۲۸\*۱۲۸ را از لینک زیر بارگذاری کنید. دقت کنید که مجموعه داده‌ی آموزشی و آزمون تفکیک شده است. برای اعتبارسنجی نیز از ۱۰٪ مجموعه داده‌ی آموزش استفاده کنید.

<http://chaladze.com/15/>

۱. با سعی و خطا تلاش کنید یک شبکه‌ی عصبی پیچشی مناسب برای دسته‌بندی تصاویر مجموعه داده‌ی مد نظر آموزش دهید.
۲. در شبکه‌ی خود اتصال(های) باقیمانده‌ی ایجاد کرده و با آموزش مجدد تغییرات احتمالی به وجود آمده را بررسی و تحلیل کنید. (صحت، سرعت همگرایی و...)
۳. در بحث بینایی کامپیوتر استخراج و تشخیص لبه<sup>۶</sup> از تصاویر به عنوان ویژگی گام مهمی تلقی می‌شود. در شبکه‌های عصبی پیچشی نیز معمولاً صفحات ویژگی<sup>۷</sup> حاصل از لایه‌ی اول Convolution می‌آموزند تا لبه استخراج کنند. با مصورسازی برای یک تصویر نمونه دلخواه به ازای خروجی‌های لایه‌ی اول از مدل خود سعی کنید مورد مذکور را نشان دهید.

انتخاب ابعاد برای کلیشه<sup>۸</sup> در شبکه‌های عصبی پیچشی با هدف استخراج ویژگی می‌تواند مشتمل بر اندازه‌های گوناگونی نظیر ۳\*۳، ۵\*۵ یا حتی ۷\*۷ باشد؛ انتخاب ابعاد کوچک‌تر معادل است با اینکه ویژگی‌های استخراجی محلی‌تر و جزئی‌تر باشد و اطلاعاتی از ویژگی‌های بزرگ تصویر در آن موجود نباشد. در دنیای واقعی، توزیع و گستره‌ی ویژگی‌های موجود در تصاویر می‌تواند در مقیاس‌های متفاوتی وجود داشته باشد و نیازمند کلیشه با ابعاد مختلف باشد. می‌توانید برای درک بهتر چالش موجود سه تصویر سگ مقابل را در نظر بگیرید که یک سبزه ثابت از کلیشه نمی‌تواند ویژگی‌های مطلوبی از هر سه تصویر استخراج نموده و اقدام به دسته‌بندی نماید؛ اگر سبزه فیلتر را بزرگ در نظر بگیریم آنگاه از تصاویری مانند سمت چپ که پیکسل‌های زیادی از سگ وجود دارد، ویژگی‌های خوبی برای دسته‌بندی استخراج می‌شود (ویژگی‌های سراسری و بزرگ)؛ اگر سبزه فیلتر را کوچک در نظر بگیریم آنگاه از تصاویری



مانند سمت راست که پیکسل‌های کمی از سگ وجود دارد، ویژگی‌های خوبی برای دسته‌بندی استخراج می‌شود. (ویژگی‌های جزئی و کوچک) برای حل مشکل مذکور شبکه‌های Inception ارائه شده است که از انواع کلیشه‌ها با ابعاد مختلف در آن استفاده می‌شود. ساختار اولین نسخه از شبکه مذکور را در تصویر زیر ملاحظه می‌کنید.

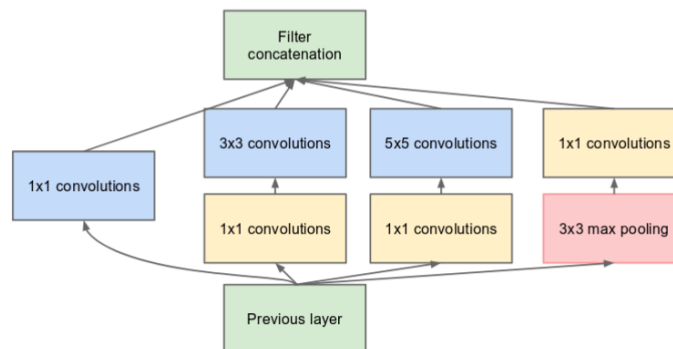
<sup>4</sup> Classifying

<sup>5</sup> Feature Extraction

<sup>6</sup> Edge

<sup>7</sup> Feature maps

<sup>8</sup> Kernel



۴. در بین مدل های قسمت قبلی، بهترین مدل را از نظر صحت پیش بینی را انتخاب کرده و تمامی لایه های Convolution با اندازه کلیشه ثابت آن را با ساختار فوق جایگزین نمایید و مدل خود را مجدد آموزش داده و نتایج حاصل را با بخش های قبل مقایسه کنید. (به طور مشابه می توانید سائز ۷\*۷ را نیز به ساختار موجود اضافه نمایید)

### انتقال یادگیری<sup>۹</sup>

آموزش مدل های شبکه ی عصبی نیازمند صرف زمان و هزینه های پردازشی می باشد و در مطالعات گوناگون رویکرد های متعددی برای اجتناب از این امر ارائه شده است. یکی از این رویکردها، استفاده از قسمت استخراج ویژگی یک شبکه ی عصبی پیچشی که قبلاً در یک مدل دیگر آموزش دیده در مدل جدید می باشد. در این حالت، قسمت دسته بند مدل قبلی از آن حذف شده و وزن های قسمت استخراج ویژگی غیرقابل آموزش<sup>۱۰</sup> تنظیم می شود؛ سپس یک دسته بند جدید به مدل افزوده شده و گام اول آموزش شروع می شود؛ در گام اول<sup>۱۱</sup> و طی چند تکرار محدود وزن های دسته بند جدید آموزش داده می شود. در گام دوم<sup>۱۲</sup>، وزن های قسمت استخراج ویژگی از انتهای به ابتدا (چرا) و طی چند مرحله قابل آموزش شده تنظیم شده و آموزش کل مدل از سر گرفته شده و مجدد چند epoch آموزش انجام می پذیرد تا مدل نهایی حاصل شود.

۵. به نظر شما چرا در گام اول انتقال یادگیری، وزن های قسمت استخراج ویژگی غیرقابل آموزش تنظیم می شود؟ پاسخ خود را نهایتاً در یک بند توضیح دهید.

حال می خواهیم با بهره از انتقال یادگیری از مدل خود که در سوال چهارم آموزش داده ایم، یک شبکه ی عصبی پیچشی برای دسته بندی مجموعه داده ی چالشی Flower<sup>۱۳</sup> پیشنهاد دهیم. مجموعه داده علاوه بر لینک زیر، در پوشه ضمیمه نیز قابل دسترس می باشد. تصاویر این مجموعه داده دارای ابعاد مختلفی هستند و از جایی که در شبکه های عصبی پیچشی مرسوم بایستی ابعاد ورودی ثابت باشد، لذا تصاویر را هم سائز و متناسب با ابعاد ورودی شبکه کنید (۱۲۸\*۱۲۸).

[https://storage.googleapis.com/download.tensorflow.org/example\\_images/flower\\_photos.tgz](https://storage.googleapis.com/download.tensorflow.org/example_images/flower_photos.tgz)

۶. از مدل حاصل در سوال چهارم برای انتقال یادگیری استفاده کرده و مجموعه تصاویر Flower را دسته بندی کنید. به نظر شما پیچیدگی دسته بند جدید چه تاثیری در روند و گام های انتقال یادگیری دارد و بایستی با چه مکانیزمی معماری دسته بند جدید طراحی شود؟

موفق باشید

<sup>۹</sup> Transfer Learning

<sup>۱۰</sup> Non-Trainable

<sup>۱۱</sup> Freezing step

<sup>۱۲</sup> Fine-tuning step