

به نام خدا
دانشکده مهندسی برق و کامپیوتر
درس یادگیری عمیق
تمرین سری دوم

هدف این سری از تمرین آشنایی با شبکه‌های کانولوشنی (CNN) و کاربرد آن‌ها در سگمنتیشن تصاویر و همچنین رگرسیون است. این سری از تمرین از دو سوال تشکیل شده است. در این سوال مجاز به استفاده از کتابخانه‌ی pyTorch می‌باشید.

Pose Estimation

1- در این سوال می‌خواهیم با دادن تصویر به یک شبکه‌ی کانولوشنی، محل سر، تنه و همچنین مفاصل دست‌ها و پاها را تعیین کنیم. بدین منظور از نسخه‌ی ساده دیتاست LSP استفاده می‌کنیم که از 2000 تصویر تشکیل شده و محل نقاط هدف به صورت یک فایل ماتریس متلب در آن داده شده است. این دیتاست از لینک زیر قابل دریافت است:

https://sam.johnson.io/research/lsp_dataset.zip

توجه داشته باشید که فایل ماتریس متلب joints.mat را می‌توان به کمک کتابخانه‌های موجود در پایتون خواند. همچنین در این تمرین نیازی به استفاده از محتویات پوشه‌ی visualized نیست. 70 درصد داده‌ها را برای آموزش، 15 درصد برای ولیدیشن و 15 درصد را برای تست به کار بگیرید. در شکل زیر نمونه‌ای از تصاویر موجود در این دیتاست مشاهده می‌شود:



شکل 1 - نمونه‌ای از دیتاست LSP

شبکه‌ی معرفی شده در مقاله‌ی زیر برای این سوال قابل استفاده است:

<https://arxiv.org/abs/1312.4659>

الف) شبکه‌ی معرفی شده در مقاله از دو قسمت Initial stage و Stage s تشکیل شده است با این هدف که در شبکه‌ی اول یک تخمین اولیه از مکان مفصل‌ها زده می‌شود و هدف شبکه‌های بعدی دقیق‌تر کردن

این تخمین است. در این قسمت شما باید ضمن خواندن مقاله و آماده‌سازی داده‌ها و برجسب‌های آن‌ها، شبکه‌ی Initial stage را پیاده‌سازی کنید و منحنی‌های loss آموزش، ولیدیشن و تست و همچنین مقادیر معیارهای PCP و PDJ را برای هر 14 مفصل رسم کنید. همچنین به عنوان نمونه، مکان مفاصل را مانند شکل (1) برای 4 تصویر از داده‌های تست رسم کنید.

ب) داده‌های قسمت آموزش را با احتمال‌های مختلف آگمنت کنید. آگمنتیشن می‌تواند شامل rotation، translation، shearing، scaling در هر دو جهت X و Y و همچنین تغییر میزان روشنایی و کنتراست باشد. توجه داشته باشید که طبیعتاً مقادیر موجود در برجسب‌ها که شامل مکان مفاصل می‌شود نیز تغییر می‌کنند. می‌توانید از کتابخانه‌ی imgaug در پایتون استفاده کنید. در نهایت موارد گفته شده در قسمت الف را مجدداً بدست آورده و مقایسه کنید.

پ) (امتیازی) شبکه مربوط به stage های بعدی را جهت دقیق‌تر کردن تخمین پیاده‌سازی کنید. استفاده از یک stage اضافه کفایت می‌کند. نتایج بدست آمده را با قسمت الف مقایسه کنید و با رسم مکان مفاصل برای 4 نمونه تست، تغییر بدست آمده را نمایش دهید (مانند شکل 6 مقاله).

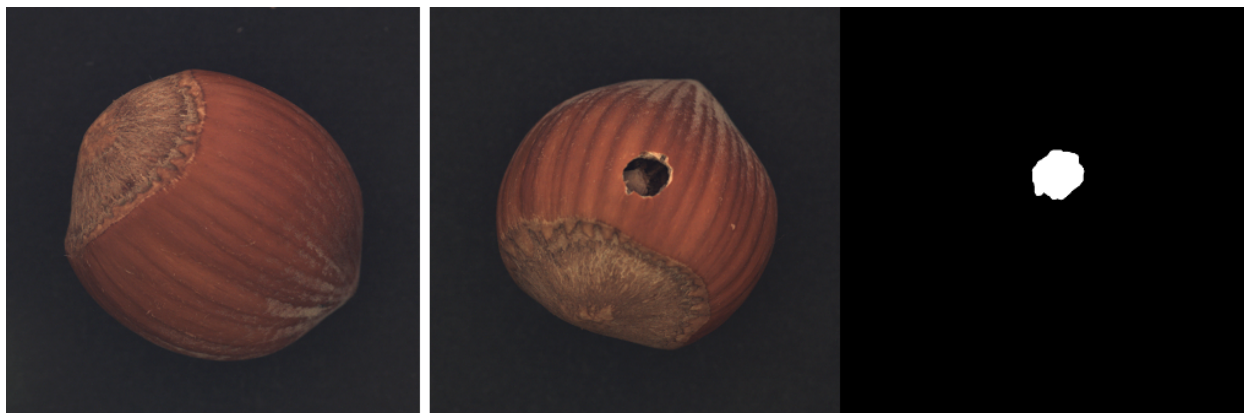
Anomaly Detection:

هدف از این بخش سوال استفاده از یک شبکه اتوانکودر برای تشخیص نقاط ناهنجاری (Anomaly) در تصاویر است. بدین منظور یکی از کلاس‌های دیتاست MVTech-AD استفاده خواهیم کرد. که می‌توانید آن را از لینک زیر دانلود نمایید.

ftp://guest:GU.205dldo@ftp.softrionics.ch/mvtec_anomaly_detection/hazelnut.tar.xz

در این تمرین از کلاس فندق در این دیتاست استفاده می‌کنیم. دیتاست از سه پوشه train، test و ground_truth تشکیل شده است.

پوشه train در دیتاست مجموعه‌ای از عکس‌های بدون عیب است. که فرآیند آموزش روی آنها صورت می‌گیرد. در پوشه test علاوه بر داده‌های بدون عیب (good)، داده‌های دیگری وجود دارد که هر یک به نحوی دارای ناهنجاری هستند. به عنوان مثال داده‌های دارای ترک خوردگی در پوشه crack آورده شده‌اند. به ازای هر داده تست دارای عیب، یک تصویر معادل در پوشه ground_truth وجود دارد که پیکسل‌های دارای آنومالی در آن را نشان می‌دهد. تصویر زیر یک نمونه از فندق سالم، یک فندق دارای سوراخ و تصویر ground_truth معادل آن را نمایش می‌دهد.



در این بخش از تمرین شما باید یک شبکه اتوانکودر را با معماری جدول-۱ آموزش دهید. در فرآیند آموزش تنها باید از داده‌های پوشه train دیتاست استفاده نمایید.

پس از اتمام آموزش، فرآیند تست بدین صورت انجام می‌پذیرد.

1. داده‌های تست را به شبکه اعمال کنید.

2. تفاضل داده ورودی و خروجی اتوانکودر، را به عنوان خطا تصویر در نظر بگیرید.

3. با پیدا کردن یک مقدار آستانه مناسب تصویر را باینری کنید.

4. با مقایسه تصویر باینری حاصل و تصویر ground_truth، accuracy شبکه را گزارش نمایید.

Layer	Outsize	Kernel_Size	Stride	Padding
input	128*128*3			
Conv1	64*64*32	4*4	2	1
Conv2	32*32*32	4*4	2	1
Conv3	32*32*32	3*3	1	1
Conv4	16*16*64	4*4	2	1
Conv5	16*16*64	3*3	1	1
Conv6	8*8*128	4*4	2	1
Conv7	8*8*64	3*3	1	1
Conv8	8*8*32	3*3	1	1
Conv9	1*1*100	8*8	1	0

Optimizer	adam
Learning Rate	*
Batch_size	32
Loss Function	MSE Loss

سوالات:

در تمامی سوالات شبکه را حداقل ۲۰۰ ایپوک آموزش داده و نمودار تغییرات خطا و دقت نهایی بر روی داده‌های تست را گزارش کنید. و از هر نوع از داده‌های معیوب حداقل یک تصویر از خروجی شبکه در گزارش بیاورید.

1. شبکه با معماری جدول-۱ پیاده سازی و آموزش دهید. شبکه‌های دیکدر و انکودر را متقارن در نظر بگیرید. برای تابع فعال ساز لایه آخر از Sigmoid و برای سایر لایه‌های از Leaky ReLU استفاده نمایید. (هم در انکودر و هم دیکدر)
2. پس از تمامی لایه‌های کانولوشن (هم در دیکدر و هم انکودر) یک لایه batch_normalization اضافه کرده و نتیجه را با سوال ۱ مقایسه نمایید.
3. شبکه را به گونه‌ای تغییر دهید تا بتواند ورودی به ابعاد $3 \times 256 \times 256$ را بپذیرد و دقت نهایی شبکه را مقایسه کنید. (راهنمایی: می‌توانید یک لایه به ابتدا انکودر و انتهای دیکدر اضافه نمایید)
4. با استفاده از تبدیل rotation داده‌های آموزش را به دو برابر افزایش داده (data augmentation) و سپس با بهترین ساختاری که از مقایسه سوال‌های قبل بدست آورده‌اید شبکه را مجددا آموزش دهید.

نکات:

- در صورت مشاهده هر گونه مشابهت کد بین هر دو دانشجو، نمره تمرین هر دو دانشجو صفر لحاظ خواهد شد.
- در صورت مشاهده هر گونه مشابهت کد با کد های موجود در صفحات اینترنتی، نمره تمرین صفر لحاظ خواهد شد. اگر بخشی از کد را از کد آماده اینترنتی استفاده می کنید که جزو قسمت های اصلی تمرین نمی باشد، حتما باید لینک آن در گزارش و کد ارجاع داده شود.
- توجه نمایید که نیمی از نمره تمرین مربوط به گزارش می‌باشد. لازم به ذکر است رعایت اصول نگارشی حائز اهمیت است.
- در نوشتن گزارش، لحاظ جزییات نوشتن گزارش الزامی است. مانند موارد زیر:
 - ارجاع دادن به مطالب و اشکالی که از مقاله و وبسایت ها گرفته شده است.
 - توضیح اشکال و جداول در caption
 - نوشتن فرمول و قرار ندادن عکس مربوط به فرمول
 - ارجاع به شکل و جدول در متن گزارش

- نوشتن نتایج شبیه سازی ها به صورت جدولی و شکل (از قرار دادن عکس نتیجه اجرای کد پرهیز شود)
- درست بودن متن از نظر قواعد دستور زبانی و نگارشی
- موارد تکمیلی در فایل template توضیح داده شده اند.
- گزارش تمرین را حتما به صورت PDF و در کنار کدهای تمرین در سایت درس آپلود نمایید.
- نحوه نامگذاری به صورت studentnumber_homeworknumber.pdf می باشد.
- زبان پیاده سازی python بوده و در این تمرین تنها مجاز به استفاده از numpy برای پیاده سازی شبکه هستید.
- برای پیاده سازی می توانید از محیط colab استفاده نمایید.
- هرگونه پرسش پیرامون تمرین را با ایمیل های aliparchekan@gmail.com و navidhasanzadeh@ut.ac.ir مطرح نمایید.

موفق باشید.