



تمرین سری ۶
درس مبانی بینایی کامپیوتر

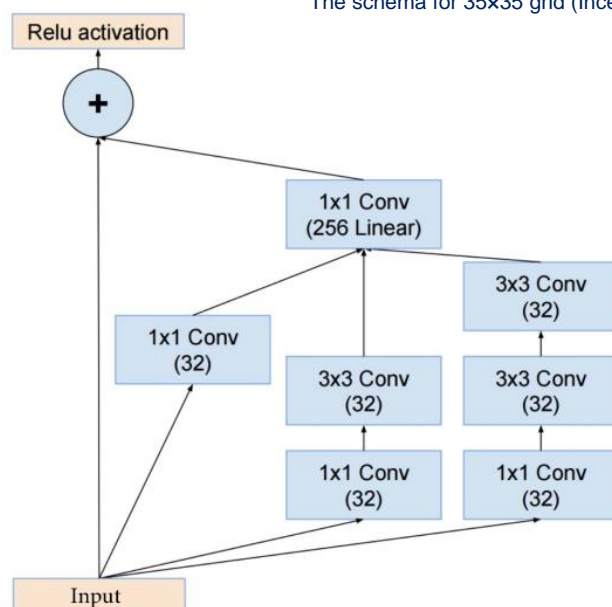
نام مدرس: دکتر محمد رضا محمدی
دستیار آموزشی مرتبط: علیرضا صدیقی مقدم، مهدی خورشیا
مهلت تحویل: ۱۴۰۲/۰۳/۲۲

۱) الف) مفاهیم ناپدید شدن گرادیان (vanishing gradient) و انفجار گرادیان (exploding gradient) و نحوه رخداد آنها را توضیح دهید. برای مطالعه می‌توانید به این لینک مراجعه کنید. (۷.۵ نمره)

ب) شبکه‌هایی که قبل از ResNet معرفی شدند برای عمیق‌تر شدن دچار کدامیک از مشکلات معرفی شده در قسمت قبل بودند و شبکه ResNet چگونه این مشکل را برطرف کرد؟ به صورت کامل توضیح دهید. (۷.۵ نمره)

۲) الف) فرض کنید که تصویر سه کاناله $n \times n$ به عنوان ورودی داریم. تعداد پارامترهای قابل آموزش و میدان تأثیر (receptive field) مربوط به ماژول زیر را محاسبه کنید (راه‌حل خود را برای تعداد پارامترها بنویسید و برای میدان تأثیر، توضیح کوتاهی از نحوه رسیدن به جواب خود ذکر کنید). (۷.۵ نمره)

The schema for 35x35 grid (Inception-ResNet-A) module of Inception-ResNet-v1 network.



ب) فرض کنید که تصویر سه کاناله $n \times n$ به عنوان ورودی داریم. دو حالت A و B را بر اساس تعداد پارامترهای قابل آموزش و میدان تأثیر (receptive field) با یکدیگر مقایسه کنید. (راه‌حل خود را برای تعداد پارامترها بنویسید و برای میدان تأثیر، توضیح کوتاهی از نحوه رسیدن به جواب خود ذکر کنید). (۷.۵ نمره)

A)

`Conv2D(filters=16, kernel_size=(3, 3), padding='valid')`

`Conv2D(filters=32, kernel_size=(3, 3), padding='valid')`

"valid" means no padding.

B)

`LocallyConnected2D(filters=16, kernel_size=(3, 3), padding='valid')`

`LocallyConnected2D(filters=32, kernel_size=(3, 3), padding='valid')`



تمرین سری ۶
درس مبانی بینایی کامپیوتر

نام مدرس: دکتر محمد رضا محمدی
دستیار آموزشی مرتبط: علیرضا صدیقی مقدم، مهدی خورشیا
مهلت تحویل: ۱۴۰۲/۰۳/۲۲

۳) نوتبوک Q3.ipynb را تکمیل کنید:

✓ الف) یک شبکه کانولوشنی ساده‌ی دلخواه بر روی دیتاست CIFAR10 آموزش دهید و نمودارهای مربوط به مقدار تابع

ضرر و دقت مدل خود را رسم و گزارش کنید. (۱۰ نمره)

✓ ب) شبکه قبلی خود را با استفاده از تکنیک داده‌افزایی (augmentation) آموزش دهید و مجدداً دو نمودار مذکور را رسم

و گزارش کنید. (۱۰ نمره)

✓ ج) نتایج و نمودارهای قسمت‌های الف و ب را با یکدیگر مقایسه و از نظر overfitting و underfitting تحلیل کنید. (۵

نمره)

✓ د) با استفاده از تکنیک انتقال یادگیری از مدل پیش آموزش دیده ResNet50 بر روی دیتاست ImageNet استفاده کنید و

نتایج آنرا با قسمت‌های قبلی مقایسه کنید. با توجه به اینکه سایز تصاویر دیتاست CIFAR10، 32×32 است و تصاویر

ImageNet 224×224 هستند در این قسمت تصاویر CIFAR10 را به اندازه 224×224 تغییر اندازه دهید. (۱۰ نمره)

امتیازی

✓ ه) در این قسمت به جای تغییر اندازه تصاویر از سه بلاک اول شبکه از پیش آموزش دیده استفاده کنید و نتایج آنرا با

قسمت‌های قبلی مقایسه کنید. به عبارت دیگر، اگر به summary مدل ذکر شده مراجعه کنید، تنها تا لایه

conv3_block4_out را در این قسمت نیاز خواهید داشت. (۱۰ نمره امتیازی)

۴) به سوالات زیر پاسخ دهید (۱۵ نمره): CV_22

✓ الف) مفهوم stride یا گام در لایه‌های کانولوشنی چیست؟ این مفهوم چه تفاوتی با Pooling دارد و چه تاثیری در

عملکرد شبکه‌های عصبی می‌گذارد؟

✓ ب) فرض کنید می‌خواهیم شبکه‌ی عصبی کانولوشنی‌ای را برای مسئله‌ی طبقه‌بندی تصاویر محصولات معیوب از سالم در

خط تولید یک کارخانه راه‌اندازی کنیم.

ب-۱) چه توابع فعال‌سازی‌ای برای لایه‌های میانی و لایه‌ی آخر شبکه‌ی خود پیشنهاد می‌کنید. دلیل خود را

بنویسید.

ب-۲) چه تابع خطایی (loss function)‌ای برای این مسئله مناسب است. دلیل خود را بنویسید

ب-۳) اگر بخواهیم دقت مدل در تشخیص محصولات معیوب از سالم زیاد باشد، به گونه‌ای که تعداد محصولات

معیوبی که به دست مشتری می‌رسد را به حداقل ممکن کاهش دهیم، عملکرد مدل را با چه معیاری بیان کنیم بهتر

است. (precision یا recall). لطفاً دلیل خود را همراه با فرضیات احتمالی خود بنویسید.

✓ ج) شبکه‌های عصبی کانولوشنی (CNN)ها در کدام یک از کاربردهای زیر موفق خواهند بود. دلیل خود را بنویسید



تمرین سری ۶ درس مبانی بینایی کامپیوتر

نام مدرس: دکتر محمد رضا محمدی
دستیار آموزشی مرتبط: علیرضا صدیقی مقدم، مهدی خورشیا
مهلت تحویل: ۱۴۰۲/۰۳/۲۲

ج-۱) طبقه‌بندی موضوع متن

ج-۲) تشخیص گوینده از روی صوت

ج-۳) تحلیل جدول مربوط به مشتریان یک فروشگاه برای پیش بینی رفتار بعدی هر مشتری، مانند جدول زیر

Name, age, address, email, purchases, browsing activity,...									
Customers	A	22	1A	a@a	1	aa	a1.a	123	aa1
	B	33	2B	b@b	2	bb	b2.b	234	bb2
	C	44	3C	c@c	3	cc	c3.c	345	cc3
	D	55	4D	d@d	4	dd	d4.d	456	dd4
	E	66	5E	e@e	5	ee	e5.e	567	ee5
	F	77	6F	f@f	6	ff	f6.f	678	ff6
	G	88	7G	g@g	7	gg	g7.g	789	gg7
	H	99	8H	h@h	8	hh	h8.h	890	hh8
	I	111	9I	i@i	9	ii	i9.i	901	ii9

د) بنظر شما شبکه‌های عصبی کانولوشنی چه مشکلاتی می‌توانند داشته باشند.

۵) فایل نوتبوک بنام ss_notebook.ipynb که به پیوست داده شده را برای انجام مرحله به مرحله‌ی تسک semantic

segmentation تکمیل کنید. در این تمرین با مفاهیم پیشرفته در کتابخانه‌ی keras برای پردازش تصویر، کار با تصاویر برای آماده سازی آن‌ها برای مسئله‌ی semantic segmentation، پیاده‌سازی شبکه‌ی عصبی برای انجام تسک semantic segmenation و نیز تحلیل آن آشنا خواهید شد. در نهایت به سوالات زیر پاسخ دهید. دیتاست را می‌توانید از این لینک دریافت کنید. (۳۰ نمره):

الف) یکی از معیارهای مهم در زمینه‌ی semantic segmentation برای ارزیابی عملکرد مدل‌های موجود IoU است. براساس فرمول زیر، یک تصویر از مجموعه‌ی داده به مدل آموزش دیده بدهید و خروجی مدل را بدست آورید، سپس میان خروجی بدست آمده و ground truth (برچسب واقعی) این معیار را محاسبه کنید. (این معیار مفهوم اصلی IoU را ندارد بلکه بصورت یک تابع loss بیان شده است که می‌تواند جایگزین BCE نیز شود و برای آموزش شبکه مورد استفاده قرار گیرد):

$$L_{iou} = 1 - \frac{\sum_{r=1}^H \sum_{c=1}^W S(r, c) G(r, c)}{\sum_{r=1}^H \sum_{c=1}^W [S(r, c) + G(r, c) - S(r, c) G(r, c)]}$$

که در آن G برچسب واقعی و S پیش بینی شبکه است و r و c سطر و ستون‌های تصاویر را نشان می‌دهند. (۵ نمره امتیازی)



تمرین سری ۶
درس مبانی بینایی کامپیوتر

نام مدرس: دکتر محمد رضا محمدی
دستیار آموزشی مرتبط: علیرضا صدیقی مقدم، مهدی خورشیا
مهلت تحویل: ۱۴۰۲/۰۳/۲۲

ب) بنظر شما تفاوت میان $BCE\ loss$ و $IoU\ loss$ در چیست؟ (۵ نمره امتیازی)

↓
Binary cross entropy

لطفا سند قوانین انجام و تحویل تمرین‌های درس را مطالعه و موارد خواسته شده را رعایت فرمایید.
موفق باشید.