



تمرین سری هشتم درس تصویرپردازی رقمی

نام مدرس: دکتر محمدی
دستیار آموزشی مرتبط: فاطمه انوری
مهلت تحویل: شنبه ۱۱ دی

۱- مقاله زیر را با دقت مطالعه کنید و به صورت خلاصه Deformable Convolution را توضیح دهید (برای توضیح بهتر از شکل‌های مقاله هم استفاده کنید).

Dai, Jifeng, et al. "[Deformable convolutional networks](#)." *Proceedings of the IEEE international conference on computer vision*. ۲۰۱۷.

۲- الف) شبکه VGG۱۹ یکی از شبکه‌های پرکاربرد برای دسته‌بندی تصویر است. جزئیات لایه‌های کانولوشنی این شبکه در شکل زیر نشان داده شده است.

Layer name	input shape			padding	stride	kernel size	Filters
Input	256	256	3				
block1_conv1	256	256	3	same	1	(3×3)	64
block1_conv2	256	256	64	same	1	(3×3)	64
block1_pool	256	256	64		2	(2×2)	
block2_conv1	128	128	64	same	1	(3×3)	128
block2_conv2	128	128	128	same	1	(3×3)	128
block2_pool	128	128	128		2	(2×2)	
block3_conv1	64	64	128	same	1	(3×3)	256
block3_conv2	64	64	256	same	1	(3×3)	256
block3_conv3	64	64	256	same	1	(3×3)	256
block3_conv4	64	64	256	same	1	(3×3)	256
block3_pool	64	64	256		2	(2×2)	
block4_conv1	32	32	256	same	1	(3×3)	512
block4_conv2	32	32	512	same	1	(3×3)	512
block4_conv3	32	32	512	same	1	(3×3)	512
block4_conv4	32	32	512	same	1	(3×3)	512
block4_pool	32	32	512		2	(2×2)	
block5_conv1	16	16	512	same	1	(3×3)	512
block5_conv2	16	16	512	same	1	(3×3)	512
block5_conv3	16	16	512	same	1	(3×3)	512
block5_conv4	16	16	512	same	1	(3×3)	512
block5_pool	16	16	512		2	(2×2)	

بعد از این لایه‌ها، برای تبدیل تنسور خروجی به بردار می‌توانیم از حالت‌های زیر استفاده کنیم و بعد از آن هم با استفاده از یک لایه کاملاً متصل دسته‌بندی را انجام دهیم. اگر مسئله دسته‌بندی ۳۰ کلاس باشد، برای هر کدام از حالت‌های زیر تعداد پارامترهای لایه مربوطه و همچنین لایه کاملاً متصل بعد از آن را محاسبه کنید.

- Flatten
- GAP
- GWAP با وزن یکسان برای تمام کانال‌ها
- GWAP با وزن متفاوت برای هر کانال
- هیستوگرام قابل آموزش با ۴ bin
- هیستوگرام قابل آموزش با ۸ bin

ب) مزایا و معایب هر کدام از حالت‌های بالا را بیان کنید.



تمرین سری هشتم
درس تصویرپردازی رقمی

نام مدرس: دکتر محمدی
دستیار آموزشی مرتبط: فاطمه انوری
مهلت تحویل: شنبه ۱۱ دی

۳- در این تمرین می‌خواهیم مسئله ناحیه‌بندی معنایی با استفاده از مجموعه داده CaltechBirds ۲۰۱۰ را شبیه‌سازی کنیم. برای انجام این تمرین از نوت‌بوک پیوست شده کمک بگیرید. در این تمرین از دو شبکه مبتنی بر معماری Encoder-Decoder استفاده خواهیم کرد. برای بخش Encoder از شبکه ResNet۵۰ استفاده کنید و بخش Decoder را به صورت متناسب اضافه کنید. در هر دو شبکه، از ورودی با ابعاد ۲۵۶ در ۲۵۶ استفاده کنید و نتایج را مقایسه کنید.

برای تکمیل این تمرین، شما نیاز است دو شبکه زیر را برای تقسیم بندی معنایی پیاده سازی کرده، و سپس با بدست آوردن نتایج، دلیل مناسب بودن یا عدم مناسب بودن هر شبکه را تحلیل نمایید.

الف) شبکه U-Net

ب) شبکه D-LinkNet

* برای انجام تمرین مجاز به استفاده از کدهای آماده نیستید و باید Decoderها را با استفاده از لایه‌های متناسب پیاده‌سازی کنید.

نکات تکمیلی:

- لطفاً پاسخ سوالات (تئوری و توضیحات پیاده‌سازی) را به طور گویا و به زبان فارسی و در صورت امکان تایپ همراه با سورس کدهای نوشته شده، فقط در یک فایل فشرده شده به شکل HW۳_YourStudentID.zip قرار داده و بارگذاری نمایید.
- منابع استفاده شده را به طور دقیق ذکر کنید.
- برای سهولت در پیاده‌سازی‌ها و منابع بیشتر، زبان پایتون پیشنهاد می‌شود. لطفاً کدهای مربوطه را در فرمت ipynb ارسال نمایید و هر کدام از موارد خواسته شده در بالا را در یک سلول جدید پیاده‌سازی نمایید.
- ارزیابی تمرین‌ها براساس صحیح بودن راه حل‌ها، گزارش مناسب، بهینه بودن کدها و کپی نبودن می‌باشد.
- در مجموع تمام تمرین‌ها، تنها ۷۲ ساعت تاخیر در ارسال پاسخ‌ها مجاز است اما پس از آن به صورت خطی از نمره شما کسر خواهد شد (معادل با روزی ۵۰ درصد).
- تمرین‌ها باید به صورت انفرادی انجام شوند و حل گروهی تمرین مجاز نیست.
- پرسش و پاسخ در رابطه با تمرین‌ها را می‌توانید در گروه مربوطه مطرح کنید.

موفق و سربلند باشید