

بسم الله الرحمن الرحيم

دانشگاه علم و صنعت ایران

پاییز ۱۳۹۹

تحويل: شنبه ۶ دی

تمرین سری هفتم

تصویرپردازی رقمی

۱. مقادیر LBP_8^1 و LBP_8^2 را برای پیکسل‌های مشخص شده از تصویر زیر محاسبه کنید. (۱۵)

0	3	5	0	2	2
3	1	2	1	0	3
1	0	3	4	2	1
2	5	3	2	0	3
5	0	4	3	1	0
1	5	2	4	1	5

برای نقاط آبی فقط LBP_8^1 و برای نقاط قرمز همه‌ی موارد ذکر شده را محاسبه کنید.

۲. در این سوال قصد داریم به دسته‌بندی تصاویر اعداد دست نویس فارسی بپردازیم. برای این منظور ابتدا از لینک زیر اقدام به دانلود مجموعه داده کنید. (۳۰)

<https://github.com/amir-saniyan/HodaDatasetReader>

برای انجام دسته‌بندی نیاز است که ویژگی‌های مناسبی از تصاویر بدست بیاورید. برای این قسمت از ویژگی‌های بافتی زیر استفاده کنید و نتایج را از لحاظ دقت و سرعت مقایسه کنید (برای کدهای آماده می‌توانید از این [لینک](#) استفاده کنید):

الف) HOG با استفاده از کد آماده

ب) LBP با استفاده از کد آماده

پ) LBP که خودتان با استفاده از دستورات پایه پیاده‌سازی می‌کنید.

سپس با استفاده از دسته‌بند [SVM](#) مدل خود را در هر حالت آموزش دهید و ارزیابی را بر روی داده‌های آزمون انجام دهید. (استفاده از دسته‌بندهای دیگر که دقت بالاتری ارائه می‌دهند نمره‌ی امتیازی دارد) نتایج ارزیابی خود را با استفاده از [Confusion Matrix](#) و [دقت](#) ارائه بدهید.

۳. به سوالات زیر پاسخ دهید: (۳۰)

الف) [learning rate](#)، [batch size](#) و [epoch](#) پارامترهای مهمی هستند که در شبکه‌های عمیق استفاده می‌شوند. با مطالعه در رابطه با این پارامترها، تاثیر هر کدام از آنها را به صورت دقیق توضیح دهید.

ب) دلایل استفاده از لایه‌های کانولوشنی (به جای لایه‌های کاملاً متصل) در کاربردهای پردازش تصویر را بیان کنید.

پ) دلایل استفاده از لایه‌ی pooling را بیان کنید.

ت) چرا MaxPooling رایج‌تر از AveragePooling است؟

ث) فرض کنید یک حجم ۳۲ در ۳۲ در ۱۰ وارد یک لایه کانولوشنی که ۱۶ فیلتر ۹ در ۹ دارد، می‌شود. مقدار گسترش مرزها چقدر باید باشد تا طول و عرض خروجی این لایه تفاوتی نکند. تعداد پارامترهای این لایه را نیز به دست آورید.

ج) اگر یک تصویر ۳ کاناله با ابعاد ۳۲ در ۳۲ وارد یک لایه کانولوشنی با ۳ فیلتر ۵ در ۵ بدون صفر افزونه و با اندازه گام ۱ شوند ابعاد خروجی چه خواهد شد؟ اگر همان تصویر را به دو لایه کانولوشنی که هر دو ۳ فیلتر ۳ در ۳ بدون صفر افزونه و اندازه گام ۱ دارند بدهیم، ابعاد خروجی چه خواهد شد؟

۴. در این سوال قصد داریم برای دسته‌بندی اعداد فارسی (همان [مجموعه داده‌ی هدا](#)) از شبکه‌های عمیق استفاده کنیم. برای معماری شبکه می‌توانید در حد نیاز از [Conv2D](#)، [MaxPooling2D](#) و [Dense](#) استفاده کنید. (معماری‌های با دقت بالاتر نمره‌ی امتیازی دارند.) برای ارزیابی این قسمت نیز از [Confusion Matrix](#) و [دقت](#) استفاده کنید و نتایج را با سوال ۲ مقایسه کنید. (۲۵)

- توجه: کدهایی که برای هر بخش نوشته‌اید را حتماً همراه با فایل گزارش در قالب یک فایل فشرده در سایت درس بارگذاری کنید. در فایل گزارش، نتایج و تصاویر بدست آمده را توضیح دهید. همچنین، تمام مراجعی (سایت، کتاب یا مقاله) که برای حل تمرین از آنها استفاده می‌کنید مرجع‌دهی کنید.