



مقدمه

هدف از این پروژه، آشنایی دانشجویان با مراحل مختلف پردازش تصویر و کاربرد آن در تشخیص کاراکترهای متنی از تصاویر (OCR) بدون استفاده از روش‌های یادگیری ماشینی است. در این پروژه، دانشجویان ابتدا مجموعه‌ای از تصاویر کپچا شامل حروف و اعداد تولید می‌کنند. سپس با بهره‌گیری از فیلترها و کرنل‌های مناسب، عملیات کاهش نویز (Denoising) و حذف تاری (Deblurring) بر روی تصاویر انجام می‌شود تا کیفیت داده‌های تصویری بهبود یابد.

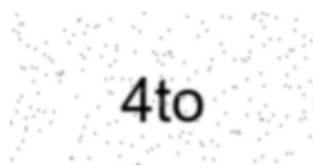
در ادامه، تصاویر به بخش‌های مجزا (Segmentation) تقسیم شده و هر بخش حاوی کاراکتر با مجموعه‌ای از الگوهای مرجع مقایسه می‌شود.

بخش اول: تولید داده

۱. در این بخش ابتدا با استفاده از قطعه کدی که در اختیارتان قرار داده شده (Captcha Generation) به تولید ۱۰ عدد تصویر رندوم ۳ کاراکتری شامل حروف کوچک انگلیسی و اعداد پردازید. (توجه کنید که در این قسمت به کتابخانه Pillow نیاز دارید)

۲. به تصاویر ایجاد شده نویز فلفل نمکی (salt and pepper noise) اضافه کنید و با تصاویر اصلی مقایسه کنید.

۳. با استفاده از کرنل‌هایی که در درس آموختید، تصویر را مات (blur) کرده و دلیل استفاده از این کرنل را توضیح دهید. در پایان این بخش خروجی شما باید شبیه تصویر زیر باشد:



بخش دوم: پیش پردازش داده‌ها

۱. میدانیم که فیلترهایی که باعث تار کردن تصویر میشوند، قابلیت از بین بردن نویز را دارند. درباره علت این قضیه در گزارشکار توضیح بدهید و از کرنلی متفاوت با قسمت ۳-۱ استفاده کنید و سعی کنید نویز تصویر را از بین ببرید.

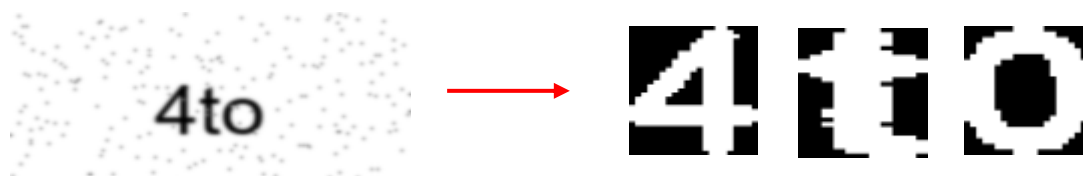
۲. در رابطه با Sharpening و De-blurring تحقیق کرده و یکی از آنها را برای از بین بردن تاری تصاویر استفاده کنید. دلیل انتخاب خود را شرح بدهید.

۳. یک روش دیگر از بین بردن نویز، باینری کردن تصاویر و حذف تکه‌های متصل به یکدیگر کوچک‌تر از آستانه (Threshold) است. این کار را انجام بدهید (انتخاب مقدار آستانه به عهده خودتان است).

بخش سوم: بخش بندی (Segmentation)

۱. خروجی قسمت قبل را گرفته و تمامی کاراکترها را به صورت تصویر باینری مجزا استخراج کنید و تعداد تصاویر استخراجی از هر تصویر را گزارش کنید.

۴. تصاویر سگمنت شده‌ی خروجی را به اندازه 64x64 پیکسل resize کنید و در پوشه ای ذخیره کنید.



بخش چهارم: Character Recognition

۵. در این مرحله، مجازید از هر روش غیر یادگیری‌محور برای سنجش شباهت بین تصاویر مرجع که در پوشه Mapset قرار داده شده و تصاویر نهایی خودتان استفاده کنید؛ از جمله روش‌هایی مانند Correlation یا سایر الگوریتم‌های مبتنی بر ویژگی‌های آماری و ساختاری تصویر. در پایان خروجی هر تصویر را به همراه یک متریک برای سنجش صحت آن در یک فایل CSV ذخیره کنید.

segment_file	predicted_label	correlation
4to_01_0.png	4	0.848607
4to_01_1.png	t	0.806126
4to_01_2.png	o	0.82111

نکات نهایی

۱. ددلاین تمرین تاریخ ۱۱ آذر ۱۴۰۴ است.

۲. تمرین دارای ارائه ۱۰ دقیقه‌ای می‌باشد و باید آمادگی ارائه آن را داشته باشید.

۳. تمرین به صورت گروه‌های دو نفره که در تمرین اول مشخص می‌شود، قابل انجام است. حتما در گزارش کار نام اعضا و شماره دانشجویی نوشته شود؛ در غیر این صورت نمره برای اسامی نوشته‌نشده تعلق نمی‌گیرد.

۴. فایل کد، گزارش کار، خروجی قسمت سگمنتیشن و فایل CSV نهایی را در پوشه‌ای با نام زیر در سامانه آپلود کنید.

CV-HW2-std#1-std#2

۵. نوشتن گزارش کار الزامی است و می‌تواند در Jupyter Notebook نیز نوشته شود.

۶. در صورت پیدا شدن موارد تقلب، نمره تمرین صفر می‌شود.

۷. شما مجاز به استفاده از کتابخانه‌های رایج پایتون مانند OpenCV هستید.