

به نام خدا



دانشگاه صنعتی امیرکبیر
دانشکده مهندسی کامپیوتر

استاد درس: دکتر صفابخش

پاییز ۱۴۰۱

درس بینائی کامپیوتر

تمرین چهارم

هدف: آشنایی با اهمیت بافت و استخراج ویژگی از آن

کد: در پیاده سازی می توانید از زبان های پایتون، متلب یا سی پلاس پلاس استفاده کنید. همچنین در تمامی موارد می توانید از کتابخانه اپن سی وی و scikit image استفاده کنید مگر اینکه صراحتاً خلاف آن در صورت سوال ذکر شده باشد.

گزارش: توجه کنید ملاک اصلی برای ارزیابی گزارش تمرین می باشد. برای این منظور گزارش را در قالب PDF تهیه کنید و برای هر سوال، تصاویر ورودی، خروجی و توضیحات مربوط به آن را ذکر کنید. همچنین اگر فرض اضافه ای در نظر می گیرید حتماً در گزارش به آن اشاره کنید.

تذکر: مطابق قوانین دانشگاه هر گونه کپی برداری و اشتراک کار دانشجویان غیرمجاز بوده و شدیداً برخورد خواهد شد. استفاده از کدها و توضیحات اینترنت به منظور یادگیری بلامانع است، اما کپی کردن غیرمجاز است.

راهنمایی: در صورت نیاز سوالات خود را می توانید در گروه مربوط به درس در تلگرام یا با ایمیل زیر مطرح کنید.

E-mail: cv.ceit.aut@gmail.com

ارسال: فایل های کد و گزارش را در قالب یک فایل فشرده با فرمت studentID_HW04.zip تا تاریخ هشت آذرماه ارسال نمایید.

تاخیر مجاز: در طول ترم، مجموعاً مجاز به حداکثر ده روز تاخیر برای ارسال تمرینات هستیید (بدون کسر نمره). این تاخیر را می توانید برحسب نیاز بین تمرینات مختلف تقسیم کنید، اما مجموع تاخیرات تمام تمرینات شما نباید بیشتر از ده روز شود. پس از استفاده از این تاخیر مجاز، هر روز تاخیر باعث کسر ده درصد نمره خواهد شد.

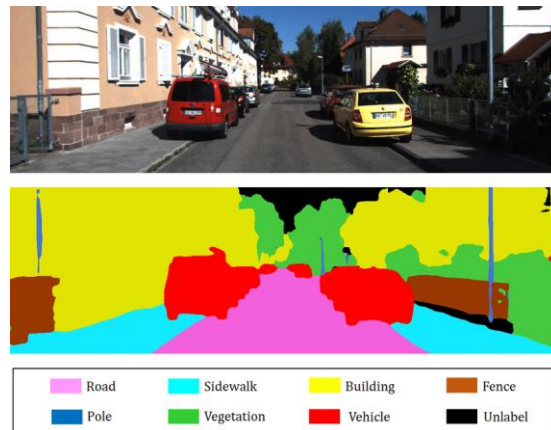
بافت به عنوان یکی از مهم ترین ویژگی های سازنده دنیای بصری، به علت اهمیت آن در درک سیستم بینایی انسان و تفکیک پذیری بالای آن برای سیستم های بینایی ماشین، از دهه ها پیش مورد توجه محققین بوده است^۱. در حقیقت، حدود نیم قرن از معرفی روش ماتریس هم رخدادی (GLCM)^۲ برای استخراج ویژگی بافت ها که در درس با آن آشنا شدید می گذرد. در سال های پس از معرفی شبکه های عصبی و یادگیری عمیق نیز، بافت یکی از ویژگی های مهمی است که این شبکه ها به آن توجه می کنند^۳. بصری سازی فیلترهای لایه های ابتدایی شبکه هایی عمیقی مانند ResNet، شباهت بالای این فیلترها را به فیلترهای گابور^۴ نشان می دهد.

¹ B. Julesz, "Visual Pattern Discrimination," in *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 84-92, February 1962, doi: 10.1109/TIT.1962.1057698.

² gray level co-occurrence matrices - R. M. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.

³ Liu, L., Chen, J., Fieguth, P. et al. From BoW to CNN: Two Decades of Texture Representation for Texture Classification. *Int J Comput Vis* 127, 74–109 (2019).

⁴ Gabor filter



شکل ۱- قطعه‌بندی معنایی: یکی از قدیمی‌ترین و مهم‌ترین وظایف سطح بالا در علم بینایی ماشین

صرف‌نظر از مزایا و معایبی^۶ که توجه تنها به بافت (بدون در نظر گرفتن شکل) برای درک تصویر ایجاد می‌کند، در اکثر کاربردهای معمول، استخراج ویژگی‌های بافت می‌تواند برای درک تصویر کافی باشد. در این تمرین با در نظر گرفتن این موضوع، قصد داریم از دو روش استخراج ویژگی‌های سطح پایین بافت، برای انجام یک وظیفه سطح بالا در بینایی ماشین استفاده کنیم.



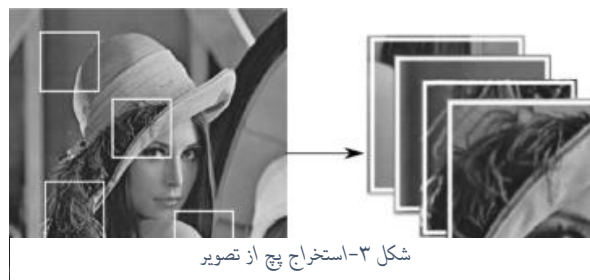
شکل ۲ - ورودی‌های تمرین

استخراج ویژگی از بافت برای قطعه‌بندی معنایی بدون نظارت^۷

هدف از قطعه‌بندی معنایی، تقسیم پیکسل‌های تصویر به گروه‌هایی است که از نظر معنایی یکسان باشند (شکل ۱). نوع بدون نظارت این وظیفه خروجی مشابهی دارد، با این تفاوت که برچسب گروه‌های مختلف در تصویر را مشخص نمی‌کند. با انجام گام‌های زیر و به کمک روش‌های استخراج ویژگی، یک مدل قطعه‌بندی معنایی توسعه دهید که تا حد ممکن تصاویر img1 و img2 را به خوبی قطعه‌بندی کند.

گام اول) استخراج ویژگی

برخلاف مسئله دسته‌بندی تصاویر، برای قطعه‌بندی معنایی نیاز به استخراج ویژگی به صورت مجزا برای هر پیکسل داریم. روش GLCM به صورت پیشفرض برای کل تصویر یک ماتریس ویژگی استخراج می‌کند؛ بنابراین



شکل ۳- استخراج پچ از تصویر

⁵ Brendel, Wieland, and Matthias Bethge. "Approximating CNNs with Bag-of-local-Features models works surprisingly well on ImageNet." *International Conference on Learning Representations*. 2018.

⁶ Geirhos, Robert, et al. "ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness." *International Conference on Learning Representations*. 2018.

⁷ Unsupervised Semantic Segmentation



شکل ۴- قطعه‌بندی تولید شده با روش شرح داده شده

نیاز است از هر پیکسل از تصویر ورودی، پچ^۸ استخراج شود (شکل ۳) و ویژگی‌ها برای هر پچ که نماینده پیکسل مرکزی آن است، محاسبه شود. به بیان ساده‌تر هر پچ را می‌توانید به عنوان یک تصویر در نظر بگیرید. پس از استخراج ماتریس GLCM از پچ‌ها، بهتر است مستقیماً آن‌ها را به ویژگی‌های آماری مانند correlation که در درس با آنها آشنا شدید تبدیل کنید.

از آنجایی که محاسبه GLCM و ویژگی‌های آماری آن نسبتاً سنگین است، می‌توانید پیش از شروع، تعداد سطوح خاکستری و ابعاد تصویر را به میزان دلخواه کاهش دهید.

برای راهنمایی بیشتر می‌توانید از منابع زیر استفاده کنید.

لینک ۱: [استخراج پچ از تصویر](#)، لینک ۲: [راهنمای GLCM](#)، لینک ۳: [پردازش موازی پچ‌ها](#)

گام دوم) خوشه‌بندی

با داشتن مجموعه‌ای از ویژگی‌ها برای هر پیکسل، می‌توان از الگوریتم‌های رایج خوشه‌بندی مانند k-means برای قطعه‌بندی تصویر استفاده کرد. پس از خوشه‌بندی پیکسل‌ها با این الگوریتم، می‌توان ساختار دو بعدی پیکسل‌ها را بازسازی کرد تا خروجی نهایی ساخته شود. یک نمونه از خروجی در شکل ۴ نمایش داده شده است.

الف) چرا استفاده از روش k-means بدون استخراج ویژگی برای قطعه‌بندی تصویرها، به ویژه در قطعه‌بندی گورخر، نتیجه خوبی نخواهد داشت؟

ب) خروجی‌های خود را با روش گفته شده برای دو تصویر پیوست، ارائه کنید.

ج) مراحل فوق را با فیلترهای گابور به عنوان استخراج‌کننده ویژگی تکرار کنید. فیلترهای گابور به ازای هر پیکسل یک مجموعه ویژگی استخراج می‌کنند؛ بنابراین نیازی به استخراج پچ‌ها از تصویر نیست.

د) خروجی دو الگوریتم را با ذکر دلایل برتری احتمالی هر یک از آن‌ها، بررسی کنید.

⁸ patch