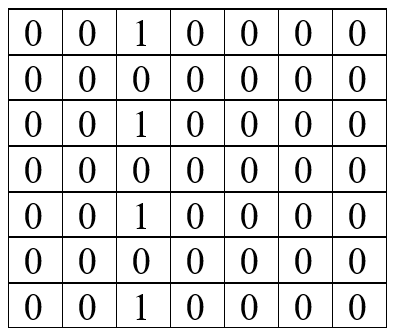
|  |  |  |
| --- | --- | --- |
|  | **به نام خدا** |  |
| **دید کامپوتری**  **پاییز 1402**  **مدرس: دکتر محمدزاده** |  | دانشگاه صنعتی شریف  دانشکده مهندسی برق |
| **تمرین اول** | **گردآورندگان: میثم امیرسرداری (تئوری) / مهرداد مرسلی (عملی)** | |

**نکات مهم پیش از انجام تمرین:**

1. تمرین‌ها پیش از تقدیم شدن به دانشجویان گرامی، پیاده‌سازی و بررسی شده‌اند. با این حال در صورت وجود هرگونه ابهام در تمرین، آن را صرفاً در گروه تلگرام مربوط به درس مطرح فرمایید.
2. هرگونه مشابهت بین تمرین‌های دانشجویان غیرمجاز است؛ لذا از کپی کردن کد از منابع اینترنتی یا سایر دانشجویان خودداری فرمایید. در صورت مشاهده، نمره آن سوال به هیچکدام از عزیزان تعلق نخواهد گرفت. توجه شود که مشابهت کدهای شما با ابزارهای مشابهت سنج بررسی خواهند شد.
3. دقت کنید کدهای شما باید قابلیت اجرای دوباره داشته باشند، در صورت دادن خطا هنگام اجرا، حتی بدلیل اشتباه تایپی، نمره صفر به آن بخش تعلق خواهد گرفت.
4. لطفا تصویری واضح از پاسخ سوالات تئوری خود آپلود کنید. در غیر اینصورت تمرین شما تصحیح نخواهد شد.
5. تمیزی و خوانایی کدهای شما نمره امتیازی سوالات عملی خواهد بود.
6. کدها، نتایج و گزارشکار بخش عملی را میتوانید به فرمت .ipynb تحویل دهید. کدهای خود را در نهایت به صورت .py نیز تبدیل کنید و هر دو فایل را بفرستید.
7. فایل های مربوط به تمرین را به فرمت HW#\_StudentNumber\_StudentName.zip آپلود کنید.

|  |
| --- |
| **سوالات تئوری 10 + 100 نمره** |

**تمرین اول: (50 نمره)**

تصویری با مقادیر شدت روشنایی به صورت نشان داده شده در شکل الف را در نظر بگیرید. ابعاد این تصویر به وسیله یک فیلتر تولیدکننده (generative kernel) 5 در 5 و با zero-padding تصویر الف برای اعمال فیلتر، به 4 در 4 کاهش پیدا کرده است. تصویر کاهش بعد یافته در شکل ب نشان داده شده است.

(ب)

(الف)

الف) فیلتر تولیدکننده را به دست آورده و به صورت یک ماتریس 5 در 5 بنویسید. (دقت نمایید که 4 پیکسل وسط شکل ب مربوط به پیکسل‌هایی از شکل الف است که به padding نیاز ندارند.) (10 نمره)

ب) فیلتر تولیدکننده قسمت الف را تجزیه کرده و به صورت یک فیلتر افقی و یک فیلتر عمودی بنویسید. فیلتر افقی را به تصویر الف اعمال کرده و نتیجه را به صورت یک ماتریس بنویسید. (6 نمره)

ج) با اعمال فیلتر قسمت الف به تصویر الف و با فرض شمارش ضرب ها ( فقط در مقادیر پیکسل های تصویر و نه مقادیر zero-pad) به چند عملیات ضرب نیاز داریم؟ با اعمال فیلتر های قسمت ب به ترتیب (اول افقی و سپس عمودی) به چند عملیات ضرب نیاز داریم؟ از مقایسه مقادیر بدست آمده چه نتیجه ای میگیرید؟ (12 نمره)

د) قسمت ج) را این بار با فرض شمارش محاسبه ضرب فقط برای مقادیر غیر صفر تکرار کنید. (10 نمره)

ه) حال فرض کنید بخواهیم با استفاده متوالی از فیلترهای افقی و عمودی به دست آمده در قسمت ب، ابعاد تصویر شکل الف را تقریبا دو برابر کنیم (13 در 13). در این قسمت فقط از فیلتر افقی به دست آمده استفاده کنید و عرض تصویر را دو برابر کنید. (نیاز به اعمال فیلتر عمودی نیست.) (12 نمره)

**تمرین دوم: (16 نمره)**

در رابطه با تخمین آستانه (thresholding) مناسب برای تولید تصویر binary از یک تصویر grayscale، الگوریتم زیر را در نظر بگیرید: (مقدار آستانه: T)

1. یک مقدار اولیه برای T در نظر بگیرید.
2. تصویر را بر اساس مقدار T به دو قسمت تقسیم کنید. پیکسل‌هایی که مقدار روشنایی بیش از آستانه دارند را گروه 1 و باقی را گروه 2 می‌نامیم.
3. مقدار میانگین و را به ترتیب برای روشنایی پیکسل‌های گروه 1 و گروه 2 محاسبه می‌کنیم.
4. مقدار جدید آستانه را بدین شکل تنظیم می‌کنیم: .
5. مراحل 2 تا 4 الگوريتم را تا زمان همگرايي تكرار می‌كنيم. (تا زمانی که اختلاف مقدار جدید آستانه با مقدار قبلی آن به قدر کافی کوچک باشد.)

حال فرض کنید برای یک تصویر فرضی، هیستوگرام روشنایی پیکسل‌ها را می‌توان به صورت مجموع دو توزیع گاوسی با مشخصه‌های و نوشت. در نظر بگیرید که بزرگتر از است و مقدار اولیه T (حد آستانه) نیز بین بیشترین و کمترین مقدار روشنایی پیکسل‌ها است. برای هر کدام از حالات زیر شرایطی را (بر اساس پارامترهای توزیع) ارائه دهید که مقدار آستانه پس از همگرایی الگوریتم چنین به دست آید:

الف) حد آستانه برابر با

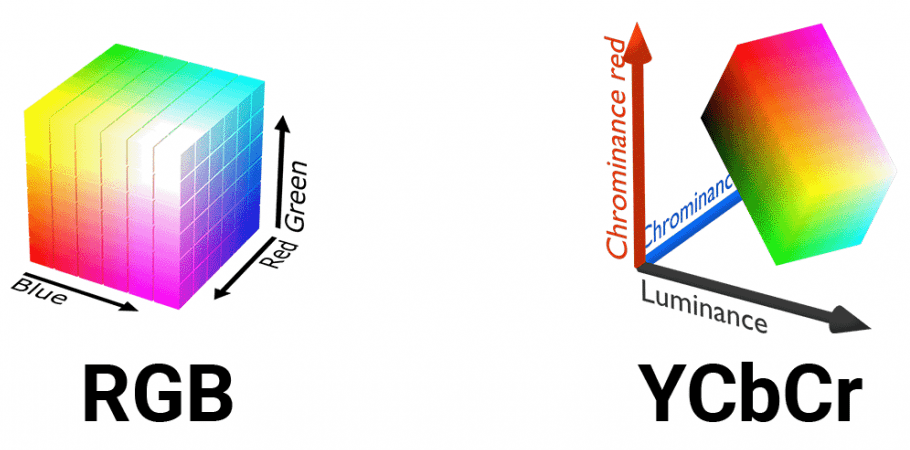
ب) حد آستانه کمتر از

اگر رخ دادن هر کدام از این حالات ناممکن است، دلیل بیاورید.

**تمرین سوم: (34 نمره)**

در کلاس درس با دو خانواده از فضاهای رنگی، HSV و RGB آشنا شدید. یک خانواده دیگر از فضاهای رنگ YCbCr است که اطلاعات رنگ و روشنایی را به سه جزء جدا می کند: Y (درخشندگی)، Cb (آبی رنگی) و Cr (قرمز رنگی). مؤلفه Y نشان دهنده روشنایی است، در حالی که مؤلفه‌های Cb و Cr به ترتیب تفاوت رنگ را از آبی و قرمز نشان می‌دهند.

شکل زیر فضای رنگی در RGB و YCbCr را مقایسه می‌کند (می‌توانید شهودی از رابطه خطی میان این دو توصیف داشته باشید):



الف) تحقیق کنید که در چه مواردی استفاده از این سیستم رنگ بر RGB ارجحیت دارد؟YCbCr در چه مواردی نسبت به RGB ضعیفتر عمل میکند؟ (6 نمره)

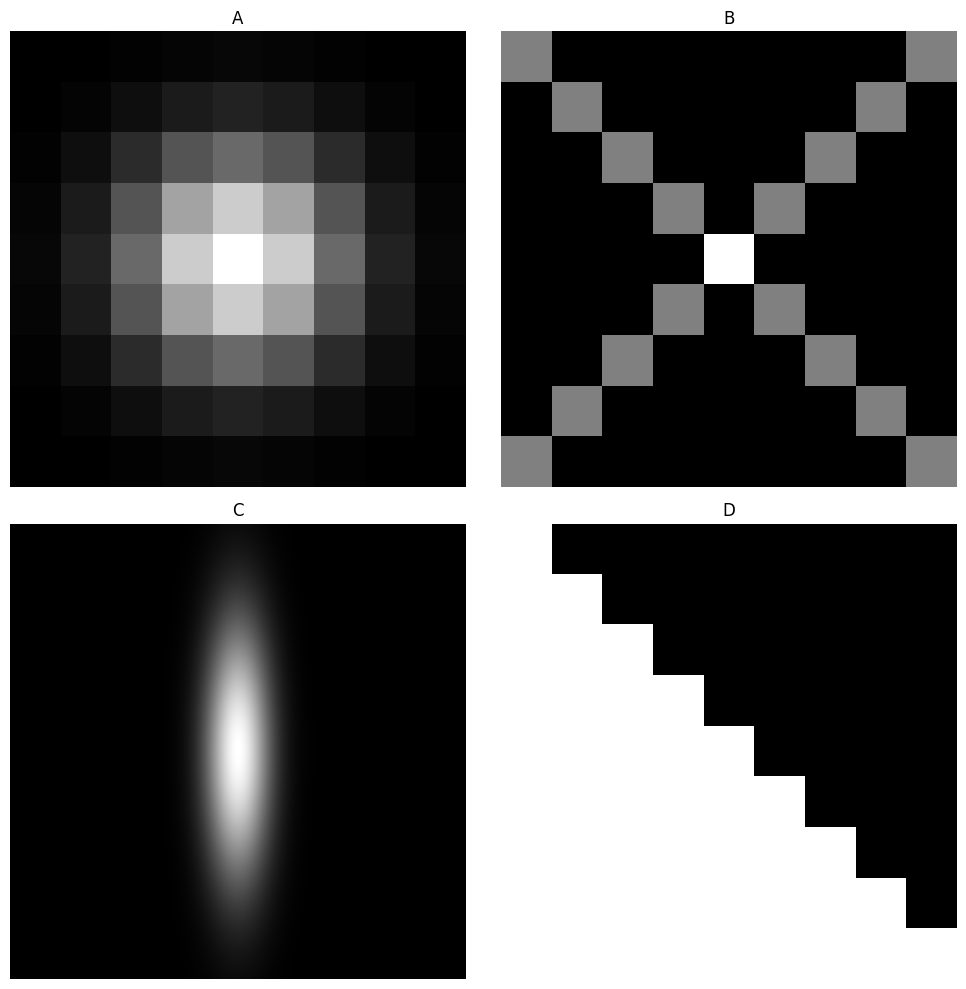
ب) تبدیل میان توصیف یک پیکسل در فضای RGB به فضای HSV و YCbCr (بر اساس استاندارد ITU-R BT.601) را جست و جو کرده و بنویسید. (10 نمره)

ج) در رابطه با نمایش YCbCr تصویر ترجیح می‌دهید Histogram Equalization را چگونه اعمال کنید که مناسبترین خروجی به دست آید؟ چرا؟ (6 نمره)

د) اگر عملیات Histogram Equalization را برای یک تصویر با سه روش مختلف، یک بار با اعمال آن بر توصیف RGB، یک بار بر توصیف HSV و در نهایت یک بار با روشی که در بخش ج پیشنهاد دادید بر توصیف YCbCr تصویر اعمال کنیم انتظار دارید سه خروجی به دست آمده چه تفاوت‌هایی با یکدیگر داشته باشند؟ کدام روش را ترجیح می‌دهید؟ (12 نمره)

**تمرین چهارم: (امتیازی) (10 نمره)**

کدام یک از فیلترهای زیر تفکیک‌پذیر هستند؟ (یعنی می‌توان آن‌ها را از ضرب فیلترهای افقی و عمودی یک بعدی به دست آورد). چرا؟



|  |  |
| --- | --- |
| **سوالات عملی** | **10 + 100 نمره** |

**تمرین اول: (40 نمره)**

1. الگوریتم Histogram Equalization تدریس شده در کلاس را با کمک کتابخانه numpy پیاده سازی کنید.( در این پیاده سازی استفاده از تابع OpenCV غیر مجاز است.) سپس الگوریتم خود را روی عکس‌های پوشه Q1 پس از تبدیل به grayscale اعمال کنید. تصاویر و histogram های آن ها را قبل و بعد از Histogram Equalization رسم کنید. ( histogram های قبل و بعد هر تصویر بر روی یک figure نمایش داده شوند. (26 نمره)
2. توضیح دهید:

الف. با توجه به تغییراتی که در سوژه‌های موجود در عکس‌ اول رخ داده است، توضیح دهید که چگونه Histogram Equalization به بازیابی قسمت‌های از دست رفته عکس کمک می‌کند. (7 نمره)

ب. با توجه به تغییراتی که در سوژه‌های موجود در عکس‌ دوم رخ داده است، توضیح دهید که چگونه Histogram Equalization به تشخیص اثر اصلی از کپی کمک می‌کند. (7 نمره)

**تمرین دوم: (30 نمره)**

هدف این تمرین این است که به کمک Joint Histogram تشخیص دهیم یک تصویر شامل چهره انسان هست یا خیر. برای این کار از یک طبقه‌بند SVC (Support Vector Machine Classifier)[[1]](#footnote-1) استفاده خواهیم کرد. ابتدا چهار مرحله زیر را طی کنید:

1. پس از تولید Joint Histogram برای تصاویر موجود در پوشه Q2/Train، آنها را در آرایه‌ای به نام x\_train ذخیره کنید. برای تولید Joint Histogram، هیستوگرام هر کانال را جداگانه حساب کرده و سه هیستوگرام را با یکدیگر concatenate کنید تا یک بردار به ازای هر عکس به دست آید.
2. برچسب‌های متناظر با هرکدام از تصاویر را از فایل TrainGroundTruth.txt موجود در پوشه Q2 خوانده و در آرایه‌ای به نام y\_train ذخیره کنید.
3. طبقه‌بند SVC را با آرایه‌های تولید شده در مرحله A و B با کد زیر آموزش دهید.

clf = SVC(C=50, kernel='poly', degree=4, random\_state=42)

clf.fit(x\_train, y\_train)

1. مراحل A و B را برای داده‌های تست نیز تکرار کنید و به کمک طبقه‌بند، برچسب داده‌های تست را پیش‌بینی کنید.

(12 نمره)

اکنون به سوالات زیر پاسخ دهید.

1. دقت پیش‌بینی چند درصد است؟ (4 نمره)
2. شماره عکس‌هایی که اشتباه پیش‌بینی شده‌اند را ذکر کنید. (2 نمره)
3. با توجه به نتایج پیش‌بینی به سوالات زیر پاسخ دهید و شرح دهید آیا نتایج، با آنچه از تئوری دید کامپیوتری انتظار داشتید مطابقت دارند یا خیر؟ ( 12 نمره)

الف. آیا طبقه‌بند به خوبی می‌تواند افراد دارای رنگ پوست متفاوت را تشخیص دهد؟

ب. آیا طبقه‌بند به خوبی می‌تواند افراد دارای ویژگی‌های خاص چهره (نظیر ریش و غیره) را تشخیص دهد؟

ج. آیا تغییر زاویه صورت بر موفقیت طبقه بند موثر است؟

**تمرین سوم: (30 نمره)**

1. پس از انتخاب بهترین لبه‌یاب (Edge Detector) با نظر خود، به کمک OpenCV تصویر صورت خود را از وبکم خوانده، آن را لبه‌یابی کرده و به صورت real-time بر روی صفحه نمایش دهید و همزمان یک ویدئو 5 ثانیه ای با فریم ریت 30 از خروجی خود ذخیره کنید. عملیات نمایش تصویر لبه‌یابی شده و ذخیره سازی در هر لحظه قبل 5 ثانیه باید با فشردن دکمه q روی keyboard متوقف و ویدیو ذخیره شود.(14 نمره)
2. برای عکس‌های موجود در پوشه Q3، لبه‌یابی را وقتی عکس ورودی در یکی از چهار حالت RGB، Gray، HSV، Binary است، بررسی کنید. برای این کار، هر تصویر موجود در پوشه Q3 را چهار بار خوانده و هر بار سیستم رنگ تصویر را به یکی از حالات ذکر شده تغییر داده و سپس آن را با لبه‌یاب Canny لبه‌یابی کنید. با مشاهده خروجی در هر حالت، چهار سیستم رنگ ذکر شده را بر حسب کیفیت لبه‌یابی رتبه‌بندی کنید. (10 نمره)
3. با توجه به تفاوت نتایج لبه‌یابی تصاویر موجود در پوشه Q3 توضیح دهید که اگر بخواهیم به کمک لبه‌یابی، وجود انسان در یک تصویر را تشخیص دهیم، چه ویژگی‌هایی در تصویر ممکن است باعث شوند که یک فرد را نتوانیم تشخیص دهیم. (6 نمره)

1. اطلاعات و نمونه کد مربوط به طبقه‌بند در این [آدرس](https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html) موجود است. [↑](#footnote-ref-1)