

به نام خدا



دانشگاه صنعتی شریف

دانشکده مهندسی برق

دید کامپیوتری

پاییز ۱۴۰۲

مدرس: دکتر محمدزاده

گردآورندگان: میثم امیرسرداری، فرهاد فلاح (تئوری) /

مهرداد مرسلی (عملی)

تمرین دوم

نکات مهم پیش از انجام تمرین:

۱. 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.
۲. 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.
۳. 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.
۴. لطفاً تصویری واضح از پاسخ سوالات تئوری خود آپلود کنید. در غیر اینصورت تمرین شما تصحیح نخواهد شد.
۵. تمیزی و خوانایی کدهای شما نمره امتیازی سوالات عملی خواهد بود.
۶. 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.
۷. فایل‌های مربوط به تمرین را به فرمت HW#\_StudentNumber\_StudentName.zip آپلود کنید.

۱۰ + ۱۰۰ نمره

سوالات تئوری

تمرین اول: (۲۰ نمره)

برای فیلتر LoG (Laplacian of Gaussian) رابطه زیر را در نظر بگیرید:

$$\nabla^2 G(r) = \frac{r^2 - 2\sigma^2}{2\pi\sigma^6} e^{-\frac{r^2}{2\sigma^2}}$$

الف) با ملاحظه تعاریف ارائه شده کلاس در رابطه با متغیر تصادفی گاوسی، نشان دهید که مقدار متوسط این عملگر صفر است، یعنی نشان دهید: (۱۰ نمره)

$$\int_0^{2\pi} \int_0^\infty \frac{r^2 - 2\sigma^2}{2\pi\sigma^6} e^{-\frac{r^2}{2\sigma^2}} d\theta dr = 0$$

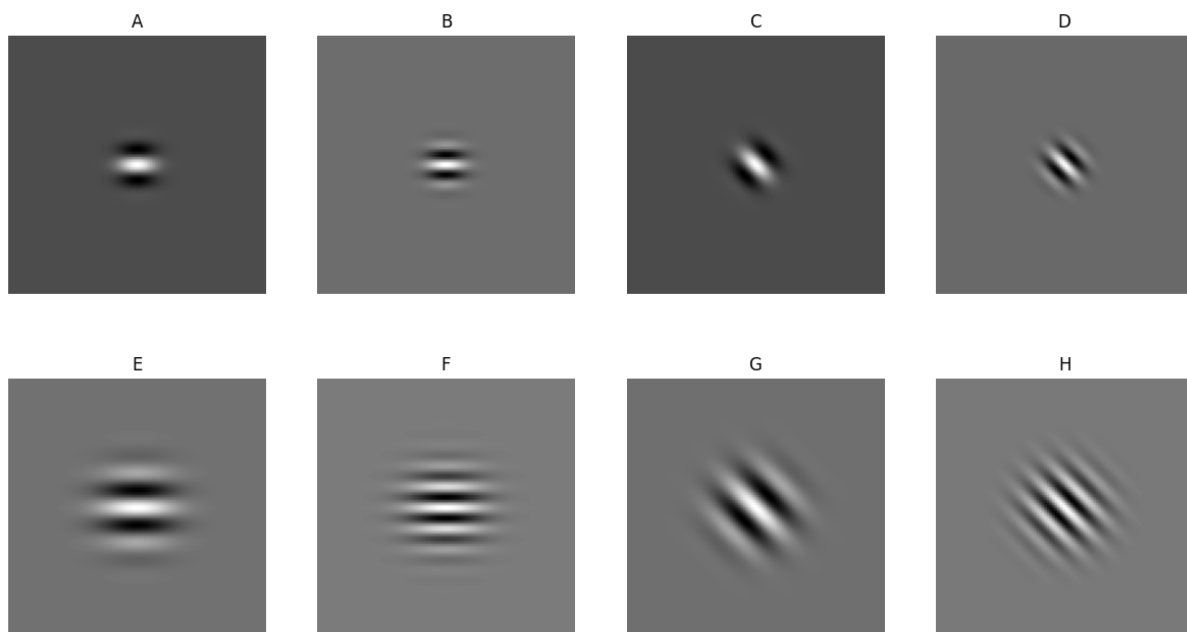
ب) اثبات کنید که مقدار متوسط هر سیگنال پیوسته کانوالو شده با این عملگر لاپلاسین نیز صفر است. (راهنمایی: در نظر داشته باشید مقدار میانگین یک سیگنال در حوزه زمان متناسب است با مقدار متناظر با مبدا در حوزه فرکانس) (۵ نمره)

ج) اگر عملگر لاپلاسین را به شکل زیر تقریب بزنیم (که یکی از تقریب‌های مرسوم در پردازش تصاویر دیجیتال است)، آیا باز هم گزاره بخش ب در صورت اعمال این کرنل به یک تصویر دیجیتال صحیح خواهد بود؟ (۵ نمره)

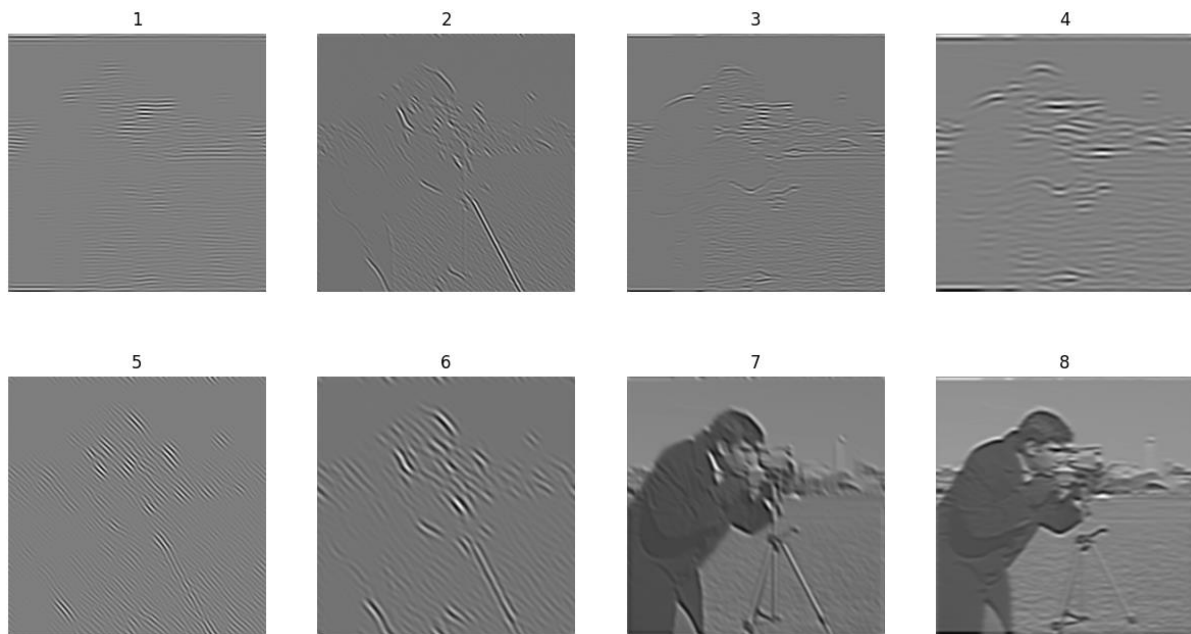
|    |    |    |
|----|----|----|
| 0  | -1 | 0  |
| -1 | 4  | -1 |
| 0  | -1 | 0  |

تمرین دوم: (۳۵ نمره)

بانک فیلتر گابور را با کرنل‌های زیر در نظر بگیرید.



الف) اگر تصاویر زیر حاصل اعمال کرنل های بالا به یک تصویر یکسان باشند، تعیین کنید هر کدام از تصاویر حاصل فیلتر کردن تصویر اصلی با کدام یک از کرنل های ذکر شده است و به اختصار توضیح دهید بر چه اساس این تناظر را برقرار می کنید؟ در حالت کلی تغییر هر کدام از پارامترهای فیلتر گابور چه تاثیری بر روی شکل کرنل و ویژگی های آشکار شده توسط آن دارد؟ (۲۰ نمره)



ب) از مفاهیم سری فوریه به یاد دارید که چگونه یک سیگنال پیوسته را می توان به صورت مجموعی از پایه های متعامد سینوسی با ضرایب مشخص توصیف و بازسازی کرد. در ادامه این سوال قصد داریم بسنجیم که آیا می توانیم به نحوی مشابه، یک تصویر را به وسیله پایه هایی از جنس فیلترهای گابور با ضرایب قابل محاسبه بازسازی کنیم یا خیر.

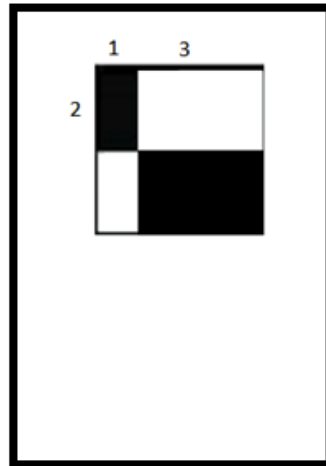
در قدم اول به نظر شما چرا چنین چیزی می تواند ممکن باشد؟ در حالت کلی فیلترهای گابور چه ویژگی هایی از تصویر را استخراج می کنند؟ (۵ نمره)

ج) آیا به نظر شما تمامی این پایه ها دو به دو به یکدیگر متعامد هستند؟ آیا شرط تعامد برای فراهم کردن یک توصیفگر کارا الزامی است؟ (برای سادگی فرض کنید منظور از پایه ها، کرنل های مختلف یک فیلتر بانک با پارامترهای متفاوت هستند) (شرط تعامد:  $\iint f_1(x,y) \overline{f_2(x,y)} dx dy = 0$  ، نیازی به حل انتگرال و اثبات ریاضی دقیق نیست.) (۵ نمره)

ه) در امتداد بحث، چنین توصیفگری از تصویر (بر پایه مجموعی از ویژگی های استخراج شده از فیلتر بانک گابور) چه مزایایی به نسبت تبدیل هایی بر پایه محاسبات صرفاً فرکانسی مانند تبدیل فوریه دو بعدی در پردازش تصویر می تواند داشته باشد؟ (۵ نمره)

### تمرین سوم: (۳۰ نمره)

فرض کنید یک تصویر باینری با مقادیر 1- و 1+ در اختیار شماست و هدف شما شمارش تعداد دفعات مشاهده الگویی مانند زیر در تصویر است. (اندازه کل الگو ۴ در ۴ پیکسل است و در آن مقادیر سیاه 1- و مقادیر سفید 1+ هستند) همچنین فرض کنید در این تصویر به غیر از این الگو با این ابعاد، الگوهای دیگری وجود ندارند.



الف) فیلتر شبه-هاری (haar-like) تعریف کنید که بتواند این الگو را در تصویر شناسایی کند. شکل دقیق فیلتر به همراه مقادیر آن را رسم کنید و آستانه مناسب برای آشکارسازی به وسیله این فیلتر را بدست آورید. (۱۵ نمره)

ب) فرض کنید تصویر انتگرال تصویر ورودی را داشته باشیم در این صورت با نوشتن رابطه متناسب برای فیلتر خود، تعداد جمع و تفریق‌های لازم برای هر بار استفاده از فیلتر را به دست آورید. (۱۰ نمره)

ج) اگر فرض کنیم اندازه تصویر ۴۰۰ در ۶۰۰ باشد برای آشکارسازی مکان این الگوها در کل تصویر به چند جمع و تفریق نیاز داریم؟ (فیلتر به تصویر بدون padding اعمال میشود). (۵ نمره)

### تمرین چهارم: (۱۵ نمره)

در شکل تصویر الف، فرض کنید خانه‌های سفید، مقادیر تصویر و مقادیر خانه‌های خاکستری، مقدار متناظر آن خانه پس از اعمال فیلتر LBP است.

|    |     |   |
|----|-----|---|
| 64 | 128 | 1 |
| 32 |     | 2 |
| 16 | 8   | 4 |

شکل ب

|     |    |    |     |
|-----|----|----|-----|
|     |    |    | $y$ |
|     | 75 | 21 |     |
| $x$ |    |    |     |

شکل الف

اگر ارزش دهی در سیستم LBP را به صورت شکل ب در نظر بگیریم، درباره علامت  $x - y$  چه میتوان گفت؟

### تمرین پنجم: (امتیازی) (۱۰ نمره)

تابع یک کرنل دو بعدی فیلتر گابور را به شکل زیر در نظر بگیرید:

$$G(x, y; \sigma, \theta, \lambda) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \cos\left(\frac{2\pi x'}{\lambda}\right)$$

که  $x' = x\cos\theta + y\sin\theta$

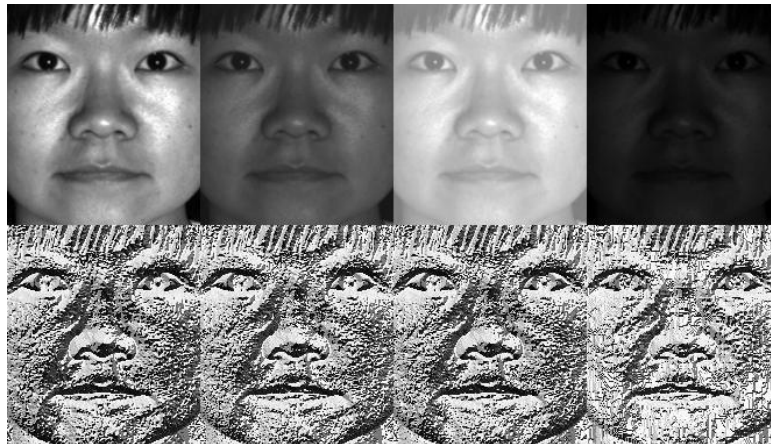
روشی برای پیاده‌سازی این فیلتر پیدا کنید که با کمترین تعداد کانولوشن فیلترهای یک بعدی قابل پیاده‌سازی باشد. رابطه این کرنل‌ها را ذکر کنید.

۱۰ + ۱۰۰ نمره

سوالات عملی

### تمرین اول: (۳۵ نمره)

در تمرین سری قبل دیدیم که وقتی از Joint Histogram برای آشکارسازی چهره انسان استفاده کردیم، حساسیت مدل به رنگ پوست افراد و نورپردازی تصویر بالا بود. یک راه حل برای رفع این مشکل، استفاده LBP برای توصیف تصویر است؛ زیرا همانطور که در شکل زیر نشان داده شده است، توصیف تصویر وقتی نورپردازی آن تغییر کند، تقریباً بدون تغییر باقی مانده است:



اکنون می‌خواهیم سیستمی طراحی کنیم که علاوه بر آشکارسازی چهره، فرد را شناسایی کرده و نام وی را اعلام کند. سیستم خود را روی دیتاست YaleFaces آزمایش می‌کنیم. این دیتاست شامل تصاویری از چهره ۱۵ فرد متفاوت است. اطلاعات مربوط به دیتاست در فایل yalefaces.csv آمده است؛ در این فایل ستون‌ها به ترتیب شامل نام سوژه، مختصات چهره سوژه در تصویر به صورت  $(x, y, w, h)$ ، نام فایل عکس و در نهایت ابعاد عکس (320x243) هستند. در این تمرین، صرفاً به ستون شماره صفر (برچسب هر عکس) و ستون شماره ۵ (آدرس عکس برای خواندن آن) احتیاج است.

برای طراحی این سیستم، مراحل زیر را طی کنید:

A. برای یافتن مختصات دقیق چهره‌ها در تصاویر، قصد داریم به جای استفاده از مختصات موجود در فایل دیتاست، آن‌ها را با استفاده از ویژگی‌های Haar بیابیم. به همین منظور کدی بنویسید که تصاویر موجود در دیتاست YaleFaces را

از پوشه Q1 به صورت Grayscale بخواند و سپس به کمک Haar Cascade Classifier موجود در پکیج OpenCV، چهره فرد را در هر عکس آشکار کند. پس از آشکارسازی چهره، طبق کادری که پیش بینی شده است، هر عکس را crop کنید. در نهایت، برای هم اندازه شدن همه نمونه‌ها، تصاویر را به سایز 130x130 تغییر دهید. برای اطمینان از درستی نتایج، یک چهره از هر فرد را به صورت رندم نمایش دهید.

B. برچسب عکس‌ها را به اعداد ۰ تا ۱۴ تبدیل کنید.

C. با تابع train\_test\_split از پکیج sklearn داده‌های train و test را ایجاد کنید؛ در این تابع، آرگومان‌های ورودی باید مشابه زیر باشند:

```
x_train, x_test, y_train, y_test = train_test_split(images, labels,
test_size=0.1, random_state=42)
```

D. با نمونه‌های آموزشی، مدل LBPHFaceRecognizer از پکیج OpenCV را آموزش دهید. این مدل بر اساس LBP چهره را شناسایی می‌کند. (۱۵ نمره)

طراحی سیستم به پایان رسید؛ اکنون به سوالات زیر پاسخ دهید:

۱. با توجه به مطالب درس، به صورت مختصر توضیح دهید Haar Cascade Classifier و LBPHFaceRecognizer چگونه کار می‌کنند. (۵ نمره)

۲. داده‌های تست را با مدل شناسایی چهره آموزش داده خود آزمایش کنید و به کمک تابع accuracy\_score از پکیج sklearn میزان دقت مدل خود در شناسایی چهره را گزارش کنید. (۵ نمره)

۳. حال تصویر image.jpg را در پوشه Q1 ملاحظه کنید. این تصویر وقتی چرخش نداشته باشد به خوبی توسط سیستم شناسایی چهره طراحی شده شناسایی می‌شود. اما وقتی چرخش در تصویر داریم، آشکارساز چهره نمی‌تواند چهره را شناسایی کند. به سیستم طراحی شده، پیش پردازشی اضافه کنید که ابتدا با Haar Cascade Classifier مخصوص تشخیص چشم از پکیج OpenCV چشم‌های سوژه را تشخیص داده و مختصات آنها در تصویر را بیابد. حال اگر مختصات گوشه چشم راست و چپ به ترتیب  $(x1, y1)$  و  $(x2, y2)$  باشند، میزان چرخش تصویر را با فرمول زیر می‌توان از مختصات چشم‌ها تخمین زد:  $\theta = \arctan(\frac{y2-y1}{x2-x1})$ . در نهایت اگر تصویر را به میزان  $\theta$  بچرخانیم، چرخش آن جبران خواهد شد. برای چرخش تصویر می‌توانید از تابع ndimage.rotate از پکیج scipy استفاده کنید. اکنون که تصویر چرخش یافت، آن را با مدل شناسایی چهره، شناسایی کرده و نام صاحب تصویر را چاپ کنید. (۱۰ نمره)

## تمرین دوم: (۴۰ نمره)

۱. الگوریتم گوشه‌یاب Harris را با استفاده از تابع موجود در OpenCV و با پارامترهای  $block\_size = 2$ ،  $ksize = 3$  و  $k = 0.04$  به تصویر موجود در پوشه Q2 اعمال کنید. سپس در اعمال Threshold به خروجی، گوشه‌هایی که امتیاز آنها از یک صدم گوشه ماکزیمم کمتر است را نادیده بگیرید. در نهایت گوشه‌های بدست آمده را با رنگ قرمز بر روی تصویر اصلی رسم کنید. (۱۰ نمره)
۲. حال می‌خواهیم این الگوریتم را مطابق شش مرحله‌ای که در درس توضیح داده شده پیاده‌سازی کنیم. مقادیر آرگومان‌ها در کد شما با مقادیر قسمت قبل یکسان و به ترتیب برابر  $block\_size = 2$ ،  $ksize = 3$  و  $k = 0.04$  هستند. همچنین مانند بخش قبل در اعمال Threshold به خروجی، گوشه‌هایی که امتیاز آنها از یک صدم گوشه ماکزیمم کمتر است را نادیده بگیرید. الگوریتم خود را به تصویر موجود در پوشه Q2 اعمال کنید و گوشه‌های بدست آمده را با رنگ قرمز نشان دهید. (۲۵ نمره)
۳. آیا گوشه‌هایی که با الگوریتم خود پیدا کردید از خروجی تابع OpenCV دقیق‌تر هستند؟ چرا؟ (راهنمایی: پاسخ شما بله است! یکی از ۶ مرحله را سخت‌گیرانه‌تر از OpenCV پیاده کنید). (۵ نمره)

## تمرین سوم: (۲۵ نمره)

درس دیدیم که برای بهبود عملکرد کرنل لبه‌یاب، مقادیر درایه‌های آن را تغییر دادیم. استدلالی در کلاس ارائه شد که نشان داد چرا فیلتر Sobel از فیلتر Prewitt بهتر است؛ آن استدلال را استدلال شماره ۱ می‌نامیم. (صفحه ۱۲ اسلایدهای لبه‌یابی) در ادامه درس، استدلالی ارائه شد که توضیح داد چرا فیلتر Gaussian از فیلتر Sobel بهتر است؛ آن استدلال را استدلال شماره ۲ می‌نامیم. (صفحه ۱۳ اسلایدهای لبه‌یابی) در زیر، کرنل‌های  $a$  تا  $d$  تقریب‌های متفاوتی از کرنل Laplacian هستند که برای لبه‌یابی مورد استفاده قرار می‌گیرند.

$$a. \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad b. \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad c. \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \quad d. \begin{bmatrix} 1 & 2 & 1 \\ 2 & -12 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

۱. با کمک استدلال ۱ و استدلال ۲ و یا سایر استدلال‌هایی که در مبحث لبه‌یابی آموخته‌اید، مشخص نمایید انتظار دارید کدامیک از کرنل‌ها ضعیف‌ترین عملکرد و کدامیک، قوی‌ترین عملکرد را در لبه‌یابی داشته باشند؟ (۱۵ نمره)
۲. کدی به زبان پایتون بنویسید که تصویر موجود در پوشه Q3 را به صورت grayscale خوانده و هریک از چهار فیلتر بالا را به آن اعمال کند. ماکزیمم و مینیمم مقادیر لبه‌های هر ۴ حالت را گزارش کنید و با مشاهده تصاویر فیلتر شده، رتبه‌بندی فیلترها را بر حسب کیفیت لبه‌یابی گزارش کنید. (تذکر: در قضاوت برای رتبه‌بندی فیلترها، فرض کنید لبه‌یابی صرفاً با اعمال فیلتر صورت می‌گیرد و عملیات تکمیلی نظیر Double Thresholding و غیره روی تصویر اعمال نخواهد شد) (۱۰ نمره)

---

<sup>۱</sup> توضیحات راجع به این آرگومان‌ها را از راهنمای OpenCV بخوانید.