

باسمه تعالی

دانشگاه صنعتی شریف

دانشکده مهندسی برق



۲۵۷۴۲ گروه ۴ - سیگنال‌ها و سیستم‌ها - بهار ۹۸-۱۳۹۷

## پروژه امتیازی

تحويل کد و گزارش این پروژه و همچنین تحويل حضوری آن در زمان تحويل حضوری فاز دوم پروژه‌ی درس انجام می‌شود.

## نحوه‌ی تحويل:

- گزارش پروژه خود را در قالب یک فایل pdf. تحويل دهید. در گزارش لازم است تمامی خروجی‌ها و نتایج نهایی، پرسش‌های متن تمرین، و توضیح مختصری از فرآیند حل مسأله‌ی خود در هر قسمت را ذکر کنید.
- کد کامل تمرین را در قالب یک فایل m. تحويل دهید. لازم است بخش‌های مختلف پروژه در sectionهای مختلف تفکیک شوند و کد تحويلی منظم و دارای کامنت‌گذاری مناسب باشد. بدیهی است تحويل دادن کدی که به درستی اجرا نشود، به منزله‌ی فاقد اعتبار بودن نتایج گزارش شده نیز می‌باشد.
- توابعی را که (در صورت لزوم) نوشته‌اید، در قالب فایل‌های m. در کنار فایل‌های گزارش و کد اصلی پروژه، ضمیمه کنید.
- مجموعه‌ی تمامی فایل‌ها (گزارش، کد اصلی، توابع، و خروجی‌های دیگر در صورت لزوم) را در قالب یک فایل zip/.rar. ذخیره کرده و تحويل دهید.
- نام‌گذاری فایل‌های تحويلی را به صورت

Opt\_Project\_StudentName\_StudentName.pdf/.m/.zip/.rar

انجام دهید.

- لطفاً آماده باشید که در زمان تحويل حضوری پروژه، کد خود را بر تصاویری با هر ابعاد داده‌شده اجرا کنید.

## نکات در مورد نمره‌ی پروژه امتیازی:

- مطابق اسلایدهای معرفی درس، این پروژه می‌تواند حداکثر به میزان ضریب مشارکت شما، که عددی کوچکتر و یا مساوی یک است (برای تعریف آن به اسلاید معرفی درس مراجعه شود)، نمره‌ی نهایی شما را افزایش دهد.
- این پروژه، به صورت گروهی و در همان گروه‌های پروژه‌ی اصلی درس انجام می‌شود و زمان تحويل آن، همان زمان تحويل حضوری فاز دوم پروژه‌ی درس است. باید در زمان تحويل فاز دوم پروژه‌ی درس، فایل‌های مربوط به این پروژه را نیز همراه داشته باشد.

شرافت انسانی ارزشی به مراتب والاتر از تعلقات دنیوی دارد. رونویسی تمارین، زیر پا گذاشتن شرافت خویشتن است؛ به کسانی که شرافتشان را زیر پا می‌گذارند به هیچ وجه نمره‌ای تعلق نمی‌گیرد.

# ۱ توضیحات اولیه

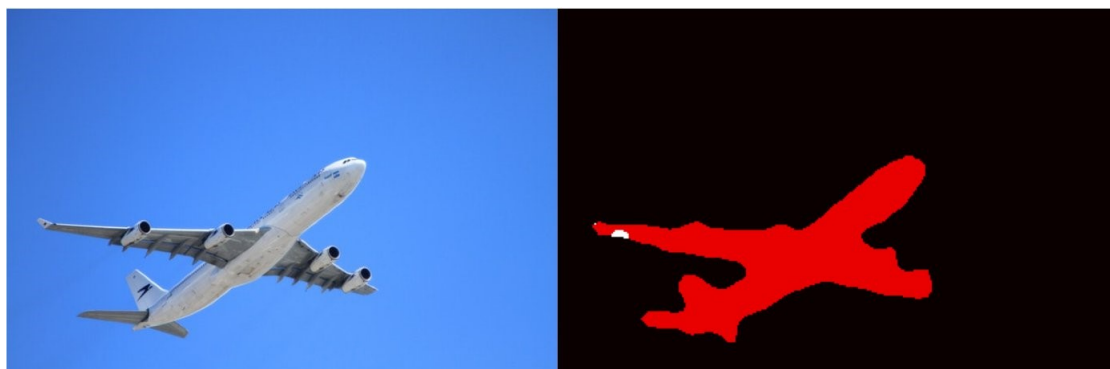
لطفاً توضیحات صفحه‌ی اول را به دقت مطالعه کنید.

در این پروژه قصد داریم به مسئله‌ی بخش‌بندی تصویر یا Image Segmentation بپردازیم که در حال حاضر از مسائل داغ حوزه‌ی بینایی ماشین است. در ابتدا دو روش برای این کار به صورت اجمالی معرفی می‌شوند. شما باید دو روش مطرح شده را پیاده‌سازی کنید و همچنین با جست و جو یک روش دیگر برای حل این مسئله ارائه کنید. در تمام طول این پروژه توجه داشته باشید که لزومی ندارد الگوریتم‌های زیر را بر روی خود تصاویر داده‌شده اجرا کنید بلکه ممکن است بعد از یک پیش‌پردازش (و کارهایی مشابه تمرین سری سوم) بتوانید به نتایج بهتری برسید!

## ۲ Image Segmentation

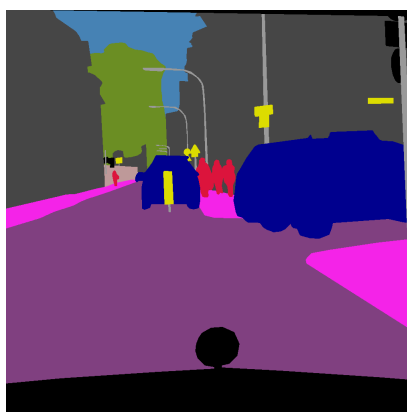
در بینایی ماشین، بخش‌بندی تصویر، به فرآیند قطعه‌بندی کردن یک تصویر دیجیتال به چند بخش (مجموعه‌ای از پیکسل‌ها) گفته می‌شود. هدف بخش‌بندی، ساده‌سازی یا تغییر در نمایش یک تصویر به چیزی است که هم معنی‌دارتر و هم برای آنالیز آسان‌تر است.

بخش‌بندی تصویر معمولاً برای پیدا کردن محل اشیاء موردنظر و مرزها (خطوط، منحنی‌ها و غیره) در تصویر استفاده می‌شود. به عبارت دقیق‌تر، بخش‌بندی تصویر، به فرآیندی گفته می‌شود که در آن، به هر پیکسل، برچسبی اختصاص داده می‌شود، به طوری که پیکسل‌هایی با برچسب یکسان، ویژگی‌های مشابهی دارند. یک مثال از بخش‌بندی، فرآیند تشخیص خیابان، پیاده‌رو، تابلوهای راهنمایی رانندگی، عابر پیاده، خودرو و ... توسط یک ماشین خودران است. یک مثال ساده از بخش‌بندی را در شکل ۱ مشاهده می‌کنید.



شکل ۱: مثالی از بخش‌بندی یک تصویر

شکل ۲: مثالی از بخش‌بندی یک خیابان است که بر خلاف شکل مثال فوق، شامل بیش از دو بخش است.



شکل ۲: مثالی از بخش‌بندی یک تصویر پیچیده

در این تمرین ما تنها به بخش‌بندی تصویر به دو بخش می‌پردازیم.

کاری که از شما می‌خواهیم ساده است! دو تصویر آسمان و هواپیما به شما داده شده است (مثل تصویر ۱) و از شما خواسته شده است که هر یک از این تصاویر را به دو بخش آسمان و هواپیما تقسیم کنید و خروجی‌ای مشابه شکل ۱ ارائه بدهید.

### ۳ روش اول: برش کمینه‌ی گرافی و الگوریتم فورد-فالکرسون

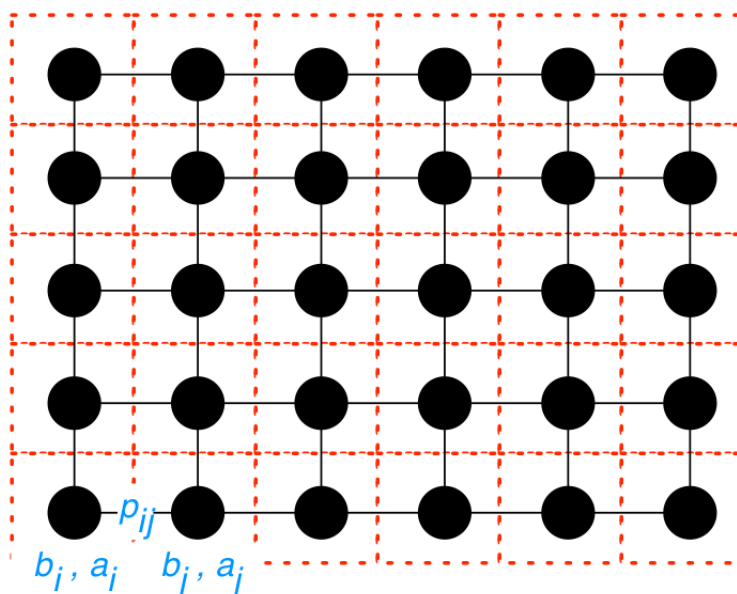
روش اولی که به صورت مختصر برای حل مسئله‌ی بخش‌بندی معرفی می‌کنیم، مبتنی بر برش کمینه‌ی گرافی و الگوریتم فورد-فالکرسون است. یک ایده این است که برای هر پیکسل  $i$  دو عدد در نظر بگیریم:

•  $a_i$  برابر احتمال اینکه پیکسل  $i$  هواپیما باشد.

•  $b_i$  برابر احتمال اینکه پیکسل  $i$  آسمان باشد.

برای هر پیکسل، جمع این دو عدد برابر یک است. برای هر پیکسل باید این دو عدد را خودتان به نحوی مناسب و معنی‌دار تعریف کنید. این تعریف می‌تواند از دانش پیشین شما در مورد آسمان و هواپیما بیاید!

می‌دانیم که اگر نقطه‌ی  $(x, y)$  در یک بخش باشد، نقاط نزدیک آن نیز احتمالاً در همان بخش هستند پس از طرفی قصد داریم کاری کنیم که پیکسل‌های نزدیک در بخش یکسان قرار بگیرند، برای این کار می‌توان عکس را به صورت یک گراف در نظر گرفت و بین پیکسل‌های مجاور یک یال قرار داد. شکل ۳ را ببینید.



شکل ۳: مثالی از گراف معادل تصویر

وزن یال بین دو پیکسل مجاور  $i$  و  $j$  نیز یک عدد مثبت است که میزان شباهت دو پیکسل را نشان می‌دهد.

می‌خواهیم در این گراف یک برش بزنییم به این معنا که راس‌ها را به دو مجموعه‌ی  $A$  و  $B$  تقسیم کنیم به این نحو که عبارت زیر کمینه شود.

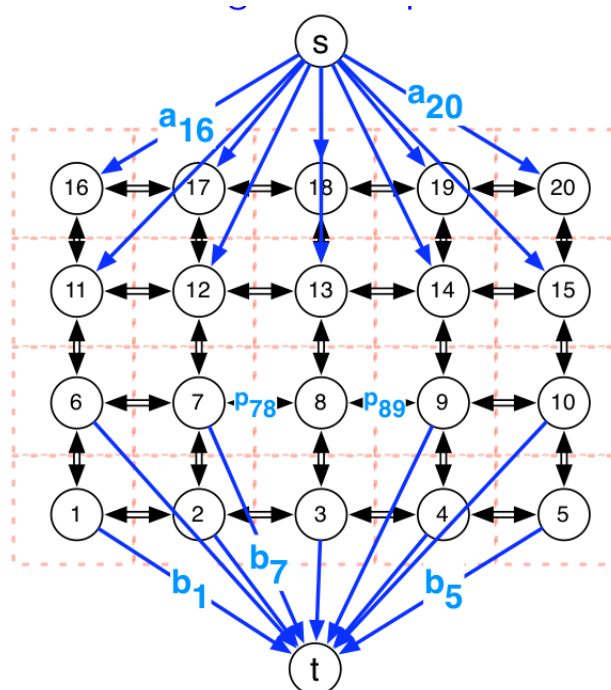
$$q(A, B) = \sum_{i \in A} a_i + \sum_{j \in B} b_j - \sum_{(i,j) \in E, i \in A, j \in B} p_{ij}$$

• چرا این هدف مطلوب است؟

- نشان دهید این کمینه‌سازی معادل است با بیشینه کردن عبارت زیر است:

$$\tilde{q}(A, B) = \sum_{i \in A} b_i + \sum_{j \in B} a_j + \sum_{(i,j) \in E, i \in A, j \in B} p_{ij}$$

- به شکل ۴ توجه کنید. نشان دهید که بیشینه‌کردن تابع فوق معادل به دست آوردن برشی کمینه در گراف این شکل است با این شرط که در این برش دو راس  $s$  و  $t$  در یک بخش نباشند. برش کمینه یعنی برشی در گراف که مجموع یال‌های بریده شده (که یک راس آن در یک سمت برش و راس دیگر در سمت دیگر برش) در آن، نسبت به تمام برش‌های ممکن دیگر کمتر باشد.
- برای حل مساله‌ی برش کمینه، می‌توان از الگوریتمی به نام Ford-Fulkerson استفاده است. بعد از مطالعه‌ی این الگوریتم، آن را پیاده‌سازی کنید و برای تصاویر داده‌شده، بخش‌بندی مناسبی پیدا کنید.
- در دستور کار خود، توضیح مختصری در مورد عملکرد الگوریتم Ford-Fulkerson بدهید. آیا این الگوریتم، الگوریتمی دقیق است و همواره برش کمینه را می‌دهد؟



شکل ۴: مثالی از گراف معادل تصویر

## ۴ روش دوم: استفاده از K-Means

الگوریتم K-Means الگوریتمی معروف برای تقسیم داده‌هایی در فضای اقلیدسی به  $k$  خوشه است. این الگوریتم قصد دارد تا با گرفتن مجموعه‌ای از نقاط مانند  $\{x_1, x_2, \dots, x_n\}$ ، که هر کدام از آنها، یک بردار هستند، آن‌ها را به  $k$  خوشه‌ی  $S = \{S_1, S_2, \dots, S_k\}$  تقسیم کند به نحوی که عبارت زیر در آن کمینه شود:

$$\sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2$$

که در آن،  $\mu_i$  برابر میانگین نقاطی است که در دسته‌ی  $i$ ام قرار گرفته‌اند. در واقع در حال حل مسئله‌ی زیر هستیم:

$$S^* = \operatorname{argmin}_{S=\{S_1, S_2, \dots, S_m\}} \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2$$

که  $S^*$  خوشه‌بندی بهینه‌ی مد نظر ماست.

این مسئله، یک مسئله‌ی بهینه‌سازی ترکیبیاتی سخت است و در واقع NP-Hard است و تا کنون هیچ الگوریتم بهینه‌ای برای حل آن یافت نشده است، بنابراین برای حل این مسئله به سمت الگوریتم‌های تقریبی و روش‌های سرانگشتی می‌رویم.

- به صورت شهودی، بیان کنید که چرا علاقه به کمینه کردن عبارت بالا داریم.
- با جست‌وجو، یک روش معروف برای به دست آوردن یک جواب معقول برای این مسئله را معرفی کنید. روش را به طور کامل توضیح دهید. توجه داشته باشید که نحوه‌ی انتخاب اولیه‌ی مرکز خوشه‌ها و همچنین شرط اتمام الگوریتم را نیز در گزارش خود ذکر کنید.
- بیان کنید که چگونه می‌توان از K-Means برای حل مسئله‌ی مطلوب ما استفاده کرد.
- با پیاده‌سازی الگوریتمی که یافته‌اید، تصاویر داده‌شده را به دو بخش تقسیم کنید.

## ۵ روش سوم: انتخابی!

در این بخش، شما باید روشی دیگر برای حل مسئله‌ی بخش‌بندی ارائه کنید. می‌توانید به الگوریتم‌های دیگر خوشه‌بندی نگاهی بیاندازید. در این بخش شما اجازه دارید تا از هر یک از توابع داخلی متلب استفاده کنید. یک ایده‌ی دیگر برای این بخش می‌تواند اعمال تغییراتی در دو روش فوق برای بهبود عملکرد آن‌ها باشد.

## ۶ مقایسه‌ی روش‌ها

به طور مختصر، سه روش فوق را با هم مقایسه کنید و مزایا و معایب هر یک را ذکر کنید. توجه کنید که این بخش از اهمیت زیادی برخوردار است پس آن را با حداکثر دقت انجام دهید.