



به نام خدا  
هوش محاسباتی (بهار ۹۹)

تکلیف شماره ۴ : بخش بندی معنایی با استفاده از خوشه بندی  
زمان تحویل: شنبه ۷ تیرماه ۱۳۹۹

بخش بندی معنایی (Semantic Segmentation) مفهومی است که در آن هر پیکسل تصویر به یک کلاس تعلق می گیرد. در واقع بخش بندی معنایی را می توان به صورت طبقه بندی تصویر در سطح پیکسل در نظر گرفت. در این مساله برخلاف مسائل طبقه بندی تصویر، خروجی یک برچسب نیست بلکه یک تصویر (با همان سائز تصویر ورودی) است که در آن هر پیکسل به یک کلاس نگاشت شده است. در این مسائل معمولاً کلاس های گوناگون به صورت رنگ مشخص می شوند. یک نمونه از تصویر ورودی و خروجی این مساله به شکل زیر است:



شکل ۱- نمونه ای از تصاویر دیتاست Cityscapes

همانطور که دیده می شود هر نمونه تصویر شامل یک تصویر اصلی از فضای شهری (سمت چپ) و یک تصویر بخش بندی شده مطلوب (سمت راست) می باشد (به تصویر سمت راست اصطلاحاً ground truth گفته می شود). در تصویر سمت راست که نتیجه حاصل از بخش بندی معنایی است، چند ماشین وجود دارد که همه جزو کلاس ماشین قرار گرفته شده و با رنگ آبی مشخص شده اند. همچنین سایر اشیا مانند درخت، علائم راهنمایی

و رانندگی و ... که در یک کلاس قرار دارند با رنگ مجزایی نشان داده شده اند. یک نمونه از کاربردهای مهم این مساله، ماشین های خودران (self-driving cars) است. در این مدل ها درک محیطی که خودرو در حال حرکت در آن هست و همچنین اشیای موجود در آن محیط مانند سایر خودروها و عابرین پیاده اهمیت بسیاری دارد. از دیگر کاربردهای بخش بندی معنایی می توان به تشخیص بیماری در تصاویر پزشکی و ویرایش تصاویر اشاره کرد. برای مثال برخی برنامه هایی که در گوشی های موبایل برای تغییر پس زمینه و ساخت استیکر با تصاویر دلخواه وجود دارد از مفهوم بخش بندی معنایی استفاده می کنند.

در این تکلیف قصد داریم یک تصویر را با استفاده از دو روش K-means و FCM بخش بندی کنیم. بدین منظور می توانید از کتابخانه های پایتون که این روش ها را پیاده سازی کرده اند استفاده نمایید و نیازی به پیاده سازی این روش ها نیست (برای این دو روش کتابخانه های متعددی وجود دارد. در این زمینه تحقیق کنید و به انتخاب خود از کتابخانه مورد نظر استفاده نمایید). ضمناً این تکلیف به صورت گروه های دو نفره می باشد. برای این تکلیف به هر گروه یک تصویر متفاوت (مشابه شکل ۱) داده می شود تا عملیات بخش بندی با استفاده از خوشه بندی های K-means و FCM را روی آن انجام دهند. به همراه صورت سوال، یک فایل فشرده بارگذاری شده که شامل تصاویر بیست که می بایست تقسیم بندی شوند. نام این تصاویر، نام اعضای گروه است لذا هر گروه باید تصویر منطبق با نام خود را برداشته و این تکلیف را روی آن پیاده سازی نماید.

به نکات زیر دقت فرمایید:

- برای انتخاب تعداد مناسب خوشه از روش elbow استفاده کنید و نمودار مربوط به تعداد خوشه و مقدار تابع هزینه (مطابق آنچه که تدریس شده است) را برای هر دو روش K-means و FCM رسم کنید. برای تعداد خوشه مقادیر ۱ تا ۱۰ خوشه را در نظر بگیرید. همچنین تصویر بخش بندی شده را یکبار بر اساس تعداد خوشه مطلوب بدست آمده از روش elbow و یکبار بر حسب تعداد خوشه های واقعی (با توجه به تعداد کلاس های تصویر ground truth) نمایش دهید. تصویر نتیجه باید به صورت یک تصویر شبیه تصویر سمت راست شکل ۱ گزارش شود. سعی کنید که از مقادیر اولیه مختلف استفاده کنید و بهترین حالت را گزارش کنید. این کار را برای هر دو روش K-means و FCM انجام داده و نتایج آن را با یکدیگر مقایسه کنید.

- برای ارزیابی نهایی باید دقت تصویر بخش بندی شده خود را محاسبه کنید. بدین منظور نیاز هست که کلاس مربوط به هر پیکسل مشخص شود. این کار را به شیوه زیر انجام دهید:

در تصویر بخش بندی شده مطلوب (ground truth) که در اختیار شما قرار داده شده است، کلاس های موجود به وسیله رنگ مشخص شده اند. بنابراین می توانید از این طریق کلاس مربوط به اشیا را استخراج کنید. برای مثال در شکل شماره ۱ رنگ آبی مربوط به کلاس "ماشین" است. فرض کنید تصویر بخش بندی شده حاصل از K-means شامل سه خوشه با شناسه های ۱ و ۲ و ۳ است. پیکسل های موجود در خوشه ۱ را در نظر بگیرید و بررسی کنید پیکسل های متناظر آنها در تصویر ground truth در کدام کلاس ها قرار دارند و کلاسی که بیشتر تعداد را دارد را به عنوان کلاس مربوط به خوشه ۱ در نظر بگیرید. برای مثال اگر خوشه ۱ شامل ۱۵۰۰ پیکسل باشد و پس از بررسی مشخص شود ۱۰۰۰ پیکسل مربوط به جاده، ۳۵۰ پیکسل مربوط به کلاس ماشین و ۱۵۰ پیکسل مربوط به درخت است، کلاس مربوط به خوشه ۱، کلاس جاده می شود. به همین شکل برای سایر خوشه ها نیز کلاس ها مشخص می شوند. پس از مشخص شدن کلاس ها و برچسب خوردن تمام پیکسل ها، می توانید دقت را به صورت زیر محاسبه نمایید:

$$accuracy = \frac{\text{تعداد پیکسل ها با برچسب درست}}{\text{تعداد کل پیکسل ها}}$$

محاسبه دقت را برای هر دو روش K-means و FCM انجام دهید.

- پیاده سازی های خود را در محیط Jupyter Notebook و در یک فایل ipynb انجام دهید. نیازی به یک فایل پی دی اف جداگانه برای توضیحات و تحلیل نتایج نیست. توضیحات خود را در همان فایل ipynb بنویسید<sup>۱</sup>. در نهایت یک فایل ipynb با عنوان شماره های دانشجویی اعضای گروه توسط سرگروه در سامانه آپلود شود.

## اختیاری:

- موارد خواسته شده در این بخش اختیاری بوده و به صورت نمره ارفاقی برای این پروژه در نظر گرفته می شود.
- از روش کاهش بعد PCA (Principal Component Analysis) برای کاهش ابعاد ویژگی های تصویر ورودی استفاده کرده و تاثیر این روش روی بخش بندی معنایی را گزارش کنید. توجه کنید که در این

---

<sup>۱</sup> با ایجاد یک سلول جدید و تغییر نوع آن از code به markdown

مساله هر پیکسل تصویر، یک داده در نظر گرفته می شود که چون تصویر RGB است دارای سه ویژگی بوده و هر پیکسل سه بعدی است. لذا کاهش بعد باید از سه بعد به دو یا یک بعد انجام شود. در این قسمت نیز برای انتخاب تعداد خوشه مناسب از روش elbow استفاده کرده و نمودار مربوط به آن را رسم کنید. تصویر بخش بندی شده را یکبار بر اساس تعداد خوشه مطلوب بدست آمده از روش elbow و یکبار بر حسب تعداد خوشه های واقعی (با توجه به تعداد کلاس های تصویر ground truth) نمایش دهید. همچنین دقت را برای هر دو روش (K-means و FCM) محاسبه کرده و تاثیر استفاده از PCA را روی نتیجه حاصل از دو روش مقایسه کنید.

- در توابع پیاده سازی شده، K-means از فاصله اقلیدسی (Euclidean distance) بین هر نمونه و مرکز خوشه مربوط به آن نمونه استفاده می کند. در این قسمت از دو معیار دیگر فاصله کسینوسی (cosine distance) و فاصله منهتن (Manhattan distance) برای خوشه بندی با استفاده از K-means استفاده کرده و مقایسه ای بین نتایج این دو فاصله و فاصله اقلیدسی انجام دهید. همانند قسمت های قبل از روش elbow برای انتخاب تعداد خوشه مناسب استفاده کرده و دقت نهایی را نیز برای این معیارها محاسبه نمایید.

(راهنمایی: این دو معیار فاصله در کتابخانه sklearn برای kmeans تعریف نشده است. می توانید از کتابخانه pyclustering استفاده کنید. در مورد نحوه استفاده از این کتابخانه مطالعه نمایید.)

- در صورتی که هر ایده ای جدید جهت بهبود عملکرد الگوریتم های مورد استفاده در این تکلیف دارید، توضیحات و نتایج آن را گزارش کنید.

موفق باشید 😊