

به نام خدا



دانشگاه تهران



دانشکده مهندسی برق و کامپیوتر

درس شبکه‌های عصبی و یادگیری عمیق

تمرین امتیازی

نام دستیار طراح	نیکو پاک‌نیا	پرسش ۱
رایانامه	nikoo.paknia@gmail.com	
نام دستیار طراح	محمد اسدزاده	پرسش ۲
رایانامه	mo.asadzadeh@ut.ac.ir	
نام دستیار طراح	محمدجواد رنجبر	پرسش ۳
رایانامه	Mohammadjavadransjbarkalahroodi@gmail.com	
نام دستیار طراح	فاطمه میرزاده	پرسش ۴
رایانامه	fatemehmirzadeh99@gmail.com	
مهلت ارسال پاسخ	۱۴۰۳.۰۳.۳۱	

فهرست مطالب

قوانین.....	۱
پرسش ۱. تولید برجسب به کمک خوشه بندی.....	۳
۱-۱. دادگان.....	۳
۲-۱. شبکه مورد استفاده.....	۳
۳-۱. آموزش شبکه.....	۴
۴-۱. ارزیابی مدل ها و مشاهده خروجی آنها.....	۴
۵-۱. خوشه بندی.....	۵
پرسش ۲. افزایش داده در مدل FaBert.....	۶
۱-۲. Data Augmentation در NLP.....	۶
۲-۲. پیش پردازش دادگان.....	۶
۳-۲. افزایش دادگان به روش Back translation.....	۶
۴-۲. تنظیم دقیق (fine-tune) مدل FaBert.....	۷
۵-۲. ارزیابی و تحلیل نتایج.....	۷
پرسش ۳. کلمه‌ی بیدار باش.....	۹
۱-۳. جمع‌آوری داده.....	۹
۲-۳. پیش‌پردازش و استخراج ویژگی.....	۱۰
۳-۳. طراحی شبکه‌ی عصبی.....	۱۰
پرسش ۴. شبکه بخش بندی تصاویر.....	۱۲
۱-۴. دادگان.....	۱۲
۲-۴. شبکه مورد استفاده.....	۱۴
۳-۴. آموزش شبکه.....	۱۴
۴-۴. ارزیابی و تحلیل نتایج.....	۱۴

شکل‌ها

- شکل ۱ مقایسه تصاویر اصلی و تولید شده توسط شبکه از دیتاست mnist ۴
- شکل ۲ مقایسه تصاویر اصلی و تولید شده توسط شبکه از دیتاست fashion mnist ۴
- شکل ۳ ارتباط کلاسترها و لیبل تصاویر ۵
- شکل ۴ کد استفاده از مترجم گوگل ۷

قبل از پاسخ دادن به پرسش‌ها، موارد زیر را با دقت مطالعه نمایید:

- از پاسخ‌های خود یک گزارش در قالبی که در صفحه‌ی درس در سامانه‌ی Elearn با نام **REPORTS_TEMPLATE.docx** قرار داده شده تهیه نمایید.
- پیشنهاد می‌شود تمرین‌ها را در قالب گروه‌های دو نفره انجام دهید. (بیش از دو نفر مجاز نیست و تحویل تک نفره نیز نمره‌ی اضافی ندارد) توجه نمایید الزامی در یکسان ماندن اعضای گروه تا انتهای ترم وجود ندارد. (یعنی، می‌توانید تمرین اول را با شخص A و تمرین دوم را با شخص B و... انجام دهید)
- **کیفیت گزارش شما در فرآیند تصحیح از اهمیت ویژه‌ای برخوردار است؛** بنابراین، لطفا تمامی نکات و فرض‌هایی را که در پیاده‌سازی‌ها و محاسبات خود در نظر می‌گیرید در گزارش ذکر کنید.
- در گزارش خود مطابق با آنچه در قالب نمونه قرار داده شده، برای شکل‌ها زیرنویس و برای جدول‌ها بالانویس در نظر بگیرید.
- الزامی به ارائه توضیح جزئیات کد در گزارش نیست، اما باید نتایج بدست آمده از آن را گزارش و تحلیل کنید.
- **تحلیل نتایج الزامی می‌باشد، حتی اگر در صورت پرسش اشاره‌ای به آن نشده باشد.**
- **دستیاران آموزشی ملزم به اجرا کردن کدهای شما نیستند؛** بنابراین، هرگونه نتیجه و یا تحلیلی که در صورت پرسش از شما خواسته شده را به طور واضح و کامل در گزارش بیاورید. در صورت عدم رعایت این مورد، بدیهی است که از نمره تمرین کسر می‌شود.
- **کدها حتما باید در قالب نوت‌بوک با پسوند ipynb تهیه شوند، در پایان کار، تمامی کد اجرا شود و خروجی هر سلول حتما در این فایل ارسالی شما ذخیره شده باشد.** بنابراین برای مثال اگر خروجی سلولی یک نمودار است که در گزارش آورده‌اید، این نمودار باید هم در گزارش هم در نوت‌بوک کدها وجود داشته باشد.
- **در صورت مشاهده‌ی تقلب امتیاز تمامی افراد شرکت‌کننده در آن، 100- لحاظ می‌شود.**
- تنها زبان برنامه نویسی مجاز **Python** است.
- استفاده از کدهای آماده برای تمرین‌ها به هیچ وجه مجاز نیست. در صورتی که دو گروه از یک منبع مشترک استفاده کنند و کدهای مشابه تحویل دهند، تقلب محسوب می‌شود.
- نحوه محاسبه تاخیر به این شکل است: پس از پایان رسیدن مهلت ارسال گزارش، حداکثر تا یک هفته امکان ارسال با تاخیر وجود دارد، پس از این یک هفته نمره آن تکلیف برای شما صفر خواهد شد.

○ سه روز اول: بدون جریمه

○ روز چهارم: ۵ درصد

○ روز پنجم: ۱۰ درصد

○ روز ششم: ۱۵ درصد

○ روز هفتم: ۲۰ درصد

- حداکثر نمره‌ای که برای هر سوال می‌توان اخذ کرد ۱۰۰ بوده و اگر مجموع بارم یک سوال بیشتر از ۱۰۰ باشد، در صورت اخذ نمره بیشتر از ۱۰۰، اعمال نخواهد شد.

○ برای مثال: اگر نمره اخذ شده از سوال ۱ برابر ۱۰۵ و نمره سوال ۲ برابر ۹۵ باشد، نمره نهایی تمرین ۹۷.۵ خواهد بود و نه ۱۰۰.

- لطفا گزارش، کدها و سایر ضمایم را به در یک پوشه با نام زیر قرار داده و آن را فشرده سازید، سپس در سامانه‌ی Elearn بارگذاری نمایید:

HW [Number] _ [Lastname] _ [StudentNumber] _ [Lastname] _ [StudentNumber].zip

(مثال: HW1_Ahmadi_810199101_Bagheri_810199102.zip)

- برای گروه‌های دو نفره، بارگذاری تمرین از جانب یکی از اعضا کافی است ولی پیشنهاد می‌شود هر دو نفر بارگذاری نمایند.

پرسش ۱. تولید برچسب به کمک خوشه بندی

در این سوال قصد داریم با یک روش self-supervised برای داده‌های تصویری برچسب تولید کنیم. در مواردی که داده بدون برچسب با حجم زیاد وجود دارد، چنین رویه‌هایی می‌تواند بسیار مفید باشد.

۱-۱. دادگان

(۵ نمره)

دو مجموعه داده mnist و fashion mnist در این سوال مورد استفاده قرار می‌گیرند. می‌توانید این دو دیتاست را از کتابخانه‌های رایج دانلود کنید.

- پس از بارگیری دیتاست‌ها، shape مجموعه آموزش و تست هر یک را بررسی کنید و یک نمونه از هر کدام را نمایش دهید.
- سپس ۲۵ درصد داده‌های آموزشی هر دیتاست را به عنوان validation set در نظر بگیرید و داده‌های هر سه مجموعه داده را نرمالایز کنید.

۱-۲. شبکه مورد استفاده

(۳۵ نمره)

- شبکه مورد استفاده یک convolutional autoencoder است. قبلاً با fully connected autoencoder آشنا شده‌اید. هم اکنون می‌خواهیم شبکه‌ای کانولوشنال را آموزش دهیم که ورودی آن تصویر و خروجی نیز همان تصویر باشد. به این ترتیب بعد از آموزش کل شبکه، بخش encoder آن را جدا کرده و به عنوان feature extractor استفاده کنیم. بخش انکدر، تصاویر را گرفته و بردار ویژگی‌ای متناسب با تصاویر را به عنوان خروجی می‌دهد. در مراحل بعد این بردار ویژگی برای خوشه بندی مورد استفاده قرار می‌گیرد.
- برای ایجاد یک convolutional autoencoder می‌توانید از لایه‌های Conv2D, Maxpooling2D, Flatten, Dense در بخش encoder آن و از لایه‌های Conv2DTranspose, ZeroPadding2D, Dense, Reshape در بخش decoder آن استفاده کنید. (توصیه می‌شود از روش functional و به کمک tensorflow مدل خود را بسازید و با لایه‌های encoder هم مدل همین بخش را تشکیل). در نهایت از autoencoder و encoder یک خلاصه به کمک متود summary() بگیرید تا از درستی

و تقارن autoencoder خود در ساختن خروجی مطمئن شوید. لایه latent برای این سوال می‌تواند بین ۴ تا ۶ بعد باشد تا نتیجه خوبی در بخش بعد بگیرید.

۱-۳. آموزش شبکه

(۱۰ نمره)

- یک مدل برای داده‌های mnist و یکی را برای fashion mnist در نظر بگیرید.
- برای بهینه‌سازی از Adam optimizer با learning rate مناسب استفاده کنید.
- معیار mse و mae را برای داده‌های آموزش و ولیدیشن به عنوان متر یادگیری در نظر بگیرید و در پایان آموزش نمودار روند آنها را رسم کنید.
- مدل‌ها را با batch size=256 و طی 10 epochs آموزش دهید.

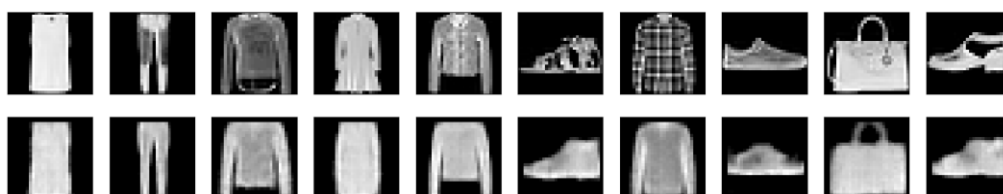
۱-۴. ارزیابی مدل‌ها و مشاهده خروجی آنها

(۱۵ نمره)

- خطای mse و mae را بر روی دادگان تست هر مجموعه گزارش کنید.
- در انتها برای هر دو دیتاست، مانند شکل زیر از هر کلاس یک تصویر اصلی و تصویر تولید شده آن توسط autoencoder مربوطه را رسم کنید.



شکل ۱ مقایسه تصاویر اصلی و تولید شده توسط شبکه از دیتاست mnist

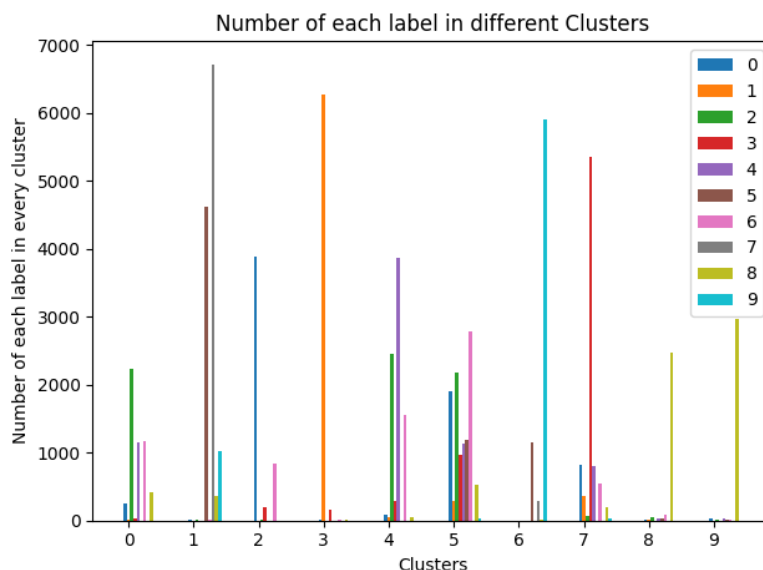


شکل ۲ مقایسه تصاویر اصلی و تولید شده توسط شبکه از دیتاست fashion mnist

۱-۵. خوشه بندی

(۳۵ نمره)

- حال به کمک encoder های مربوط به هر مجموعه داده، برداری متشکل از ویژگی‌های کلیه داده‌های train, validation, test به دست آورید. همچنین تمام برچسب‌های این سه مجموعه داده را هم به یک بردار تبدیل کنید.
- سپس به کمک الگوریتم KMeans بر روی بردار ویژگی‌ها خوشه‌بندی انجام دهید. برای خوشه‌هایی با اندازه بین ۵ تا ۱۵ این الگوریتم را اجرا کنید، سپس بر اساس معیار silhouette score بهترین تعداد خوشه را برای mnist و سپس fashion mnist را پیدا کنید. نمودار silhouette score را برای تعداد خوشه متفاوت رسم کنید و در مورد آن توضیح دهید.
- بار دیگر و بعد از پیدا کردن تعداد خوشه مناسب، الگوریتم KMeans را اجرا کنید و در نمودار مانند نمودار زیر برای هر خوشه نمودارهای ستونی رسم کنید به نحوی که تعداد یا درصد هر کدام از برچسب‌های اصلی در هر خوشه مشخص باشد و بررسی کنید که هر خوشه در حقیقت کدام برچسب را کشف کرده است. همچنین در مورد کیفیت عملکرد این روش بر هر کدام از مجموعه داده‌ها نظر دهید.



شکل ۳ ارتباط کلاسترها و لیبل تصاویر

پرسش ۲. افزایش داده در مدل FaBert

در این بخش با مفهوم افزایش داده (Data augmentation) در زمینه پردازش زبان طبیعی (NLP) آشنا می‌شوید و به پیاده سازی و تحلیل نتایج آن می‌پردازید.

۲-۱. Data Augmentation در NLP

(۱۵ نمره)

در مورد رویکردهای افزایش داده در زمینه‌ی پردازش زبان‌های طبیعی (NLP) تحقیق کنید و روش Back translation را توضیح دهید.

۲-۲. پیش پردازش دادگان

(۲۰ نمره)

مجموعه داده‌ای که با آن کار خواهید کرد DeepSentiPers می‌باشد. این مجموعه داده برای تسک Sentiment Analysis جمع آوری شده و شامل بیش از ۷۰۰۰ هزار کامنت برای محصولات عرضه شده در دیجیکالا می‌باشد. برای دسترسی به [دادگان آموزش](#) و [دادگان آزمون](#) از لینک‌های مربوط به آن استفاده کنید.

در این دیتاست هر کدام از نمونه‌ها می‌توانند برچسب ۲-، ۱-، ۰، ۱+، ۲+ داشته باشند. برچسب‌های (۱- و ۲-) و (۱+ و ۲+) را با یکدیگر ادغام کنید. و یک دیتاست سه کلاسه با کلاس‌های ۱+ و ۱- و ۰ بسازید.

اطلاعات آماری شامل توزیع کلاس‌ها و تعداد توکن‌های این مجموعه داده را ارائه دهید و در صورت نیاز پیش پردازش‌های لازم را انجام دهید.

۲-۳. افزایش دادگان به روش Back translation

(۲۰ نمره)

Back translation را برای افزایش دادگان پیاده‌سازی کنید، برای این منظور می‌توانید از کتابخانه googletrans استفاده نمایید. پس از افزایش داده اندازه دیتاست افزوده شده نسبت به دیتاست اولیه باید

```
!pip install googletrans==3.1.0a0
from googletrans import Translator
```

شکل ۴ کد استفاده از مترجم گوگل

دو برابر شده باشد.

۱۰ نمونه از دیتاست اولیه و معادل آن در دیتاست افزوده شده را نمایش داده و کیفیت نمونه‌های افزوده شده را بررسی کنید.

۲-۴. تنظیم دقیق (fine-tune) مدل FaBert

(۲۰ نمره)

[FaBert](#) یک مدل مبتنی بر Bert می‌باشد که روی دادگان فارسی آموزش دیده‌است.

دادگان آموزش مرحله قبل - شامل دیتاست ابتدایی و دیتاست افزوده شده- را بطور جداگانه روی این مدل fine-tune کنید. با توجه به اینکه حجم دادگان چندان زیاد نیست برای جلوگیری از بیش برآزش همه لایه‌های مدل FaBert را بجز لایه آخر را فریز (freeze) کنید. تعداد اپاک را ۵ و نرخ یادگیری را $10e-5$ در نظر بگیرید.

دقت و مقدار تابع هزینه روی دادگان ارزیابی (validation data) در هر اپاک را برای هر دو مدل محاسبه و ذخیره کنید.

۲-۵. ارزیابی و تحلیل نتایج

(۲۵ نمره)

با توجه به نتایج بدست آمده در قسمت قبل موارد زیر را ارائه دهید:

- نمودار اپاک-دقت برای هر یک از مدل‌های آموزش دیده روی دیتاست ابتدایی و دیتاست افزوده شده (در یک نمودار رسم شود)

- نمودار اپاک-تابع هزینه برای هر یک از مدل‌های آموزش دیده روی دیتاست ابتدایی و دیتاست افزوده شده (در یک نمودار رسم شود)

- ماتریس آشفتگی و f1-score برای هر یک از مدل‌ها پس از آموزش گزارش دهید.
- چند مورد از نمونه‌هایی دادگان ارزیابی را که مدل به اشتباه دسته بندی کرده را نمایش دهید.
- در نهایت تحلیل خود را از نتایج بدست آماده بنویسید.

پرسش ۳. کلمه‌ی بیدار باش

کلمه بیدار باش یا "wake word" یک عبارت یا کلمه است که دستگاهی که با تکنولوژی شناسایی صدا مجهز شده است را فعال می‌کند. در واقع، این کلمه یا عبارتی است که شما به آن می‌گویید تا توجه دستگاهی مانند Siri اپل، Alexa از آمازون یا Google Assistant را جلب کنید. به عنوان مثال، گفتن "سیری" یک آیفون یا دستگاه اپل دیگر را بیدار می‌کند، در حالی که "الکسا" یا "هی گوگل" همین کار را برای دستگاه‌های Echo از آمازون یا Google Home انجام می‌دهد.

اهمیت کلمات بیدار باش در این است که این‌ها قابلیت فعال کردن بازوهای صوتی یا دستگاه‌های دیگر کنترل شده توسط صدا را به شکل کارآمدی دارند. با استفاده از کلمه بیدار باش، کاربران می‌توانند با بازوهای صوتی یا دستگاه‌های کنترل شده با صدا تعامل کنند بدون اینکه نیاز به فشردن دکمه‌ای یا تعامل فیزیکی با آنها داشته باشند. این تعامل بدون دستی به ویژه در شرایطی که تعامل دستی غیر مناسب یا غیر ممکن است، مانند حین رانندگی، پخت و پز یا انجام چند وظیفه، ارزشمند است.

علاوه بر این، کلمات بیدار باش کمک می‌کنند از فعال شدن تصادفی دستگاه‌ها جلوگیری شود، با اطمینان حاصل می‌کنند که آنها فقط زمانی واکنش می‌دهند که به ویژه توسط کاربر فراخوانی شوند. بدون کلمه بیدار باش، دستگاه ممکن است به طور مداوم برای دستورات گوش داده شود، که باعث نگرانی‌های حریم خصوصی و فعال شدن‌های ناخواسته می‌شود. استفاده از کلمه بیدار باش این مسائل را با ارائه یک سیگنال روشن به دستگاه برطرف می‌کند که کاربر قصد تعامل با آن را دارد. در این تمرین قصد ساخت مدل یادگیری عمیقی داریم که قابلیت تشخیص کلمه‌ی بیدار باش شما را داشته باشد.

۳-۱. جمع‌آوری داده

(۱۰ نمره)

به صورت کلی، این مساله را می‌توان به یک مساله‌ی دسته‌بندی صوت کاهش داد، به این صورت که دو کلاس شامل ۱- کلمه‌ی بیدار باش، ۲- سایر اصوات داریم. بنابراین داده‌های اصلی ما شامل صوت‌هایی از گفته شدن کلمه‌ی بیدار باش و صداهای ضبط شده از کلمات دیگر به همراه سکوت یا صدای پس زمینه می‌باشد.

- برای این کار دو تابع record_wake_word و record_background_sound بنویسید. با استفاده از این دو تابع، هربار دقیقا دو ثانیه از صد نمونه صوت پس زمینه و کلمه‌ی بیدارباش ضبط کنید. (۱۰ نمره)
- انجام کد بخش بالا الزامی است، در صورتی که کیفیت صوت‌های ضبط شده‌ی شما به اندازه‌ی کافی نیست می‌توانید از این [لینک](#) از مجموعه داده‌های مربوط به این کار استفاده کنید. (یا هر مجموعه داده‌ی دیگری که در سطح وب پیدا کردید، با لینک در گزارش ذکر کنید).

۲-۳. پیش‌پردازش و استخراج ویژگی

(۳۰ نمره)

حال که داده‌های خود را جمع‌آوری کرده‌ایم باید آن‌ها را پیش‌پردازش کرده و ویژگی‌های مربوط به صوت را از آن استخراج کرد.

- درباره نحوه‌ی پیش‌پردازش داده‌های صوتی تحقیق کنید و آن‌ها را توضیح دهید. (۵ نمره)
- درباره ویژگی‌های مختلف صوت که قابل استخراج هست تحقیق کنید و مهم‌ترین ویژگی‌ها را نام برده و توضیح دهید. می‌توانید ویژگی‌هایی پایه‌ای از صوت را نام ببرید که در کاربردهای مختلف پردازش صدا استفاده از آن‌ها می‌تواند مضر باشد در حالی که در کاربردهای دیگر مفید است؟ (۵ نمره)
- درباره روش‌های data augmentation در صوت تحقیق کنید و تعدادی از آن‌ها را برای این بخش پیاده‌سازی کنید. (۲۰ نمره)
- پیش‌پردازش و استخراج ویژگی از صوت‌های ضبط شده را پیاده‌سازی کنید. (۱۰ نمره)

۳-۳. طراحی شبکه‌ی عصبی

(۵۰ نمره)

در اینجا قصد داریم که یک شبکه‌ی عصبی ساده پیاده‌سازی کنیم تا وظیفه دسته‌بندی این صوت‌ها را داشته باشد.

- با توجه به اینکه صوت‌های شما طول ثابتی دارند چه شبکه‌ای پیشنهاد می‌کنید؟ (۲۰ نمره)
- در صورتی که صوت‌ها طول ثابتی نداشتند شما نیاز داشتید که frame‌های مختلف از یک صوت را به عنوان ورودی دهید. در این صورت چه راه‌حلی پیشنهاد می‌کنید؟ (۱۰ نمره)
- شبکه‌ی مورد نظر خود را پیاده‌سازی کنید و عملکرد آن را روی داده‌های تست ارزیابی کنید و نمودارهای مربوط به دقت، خطا و ماتریس درهم‌ریختگی را نمایش دهید. (۳۰ نمره)

پرسش ۴. شبکه بخش بندی تصاویر

هدف از این سوال، پیاده سازی شبکه‌های مختلف جهت بخش‌بندی تصاویر می‌باشد. تمرکز این تمرین بر روی پیاده سازی دو شبکه UNet و Ta-Unet است.

UNet یک معماری شبکه عصبی است که برای بخش‌بندی تصاویر بسیار موثر است. این شبکه دارای ساختاری بازگشتی است که از لایه‌های کاهشی (down sampling) و لایه‌های افزایشی (up sampling) تشکیل شده است. این ساختار امکان استخراج ویژگی‌های تصاویر در سطوح مختلف جزئیات را فراهم می‌کند و به تشخیص دقیق‌تر اجسام کمک می‌کند.

Ta-Unet یک نوع ارتقاء یافته از UNet است که از تکنیک‌هایی مانند توجه (Attention) برای بهبود عملکرد و دقت بخش‌بندی تصاویر استفاده می‌کند. این شبکه از مکانیزم‌هایی مبتنی بر توجه برای تمرکز بر بخش‌های مهم تصاویر و رهایی از بخش‌های نامرتبط استفاده می‌کند که باعث بهبود کیفیت و دقت نتایج می‌شود.

۴-۱. دادگان

(۳۰ نمره)

مجموعه داده SUIM مجموعه داده‌ای شامل تصاویر و بخش‌بندی اجسام و موجودات زیر آب است، که این مجموعه دادگان در پوشه همراه صورت سوال برایتان ارسال شده است (لازم به ذکر است دیتاستی که خدمتتان ارسال شده متفاوت با دیتاست مقاله است ولی می‌بایست شبکه‌های مذکور در مقاله با رعایت تمامی نکات را روی دیتاست جدید پیاده‌سازی بفرمائید).

این ساختار دیتاست به صورت زیر است:

پوشه Train_val: شامل ۱۵۲۵ نمونه جفتی برای آموزش و اعتبارسنجی است. این دایرکتوری شامل زیردایرکتوری‌های تصاویر و ماسک‌ها می‌باشد. داخل دایرکتوری Train_val:

images: تصاویر RGB از صحنه‌های زیرآب برای آموزش و اعتبارسنجی.

masks: برچسب‌های بخش‌بندی متناظر با تصاویر RGB.

هر رنگ RGB در ماسک یک دسته مختلف اشیا را نمایان می‌کند.

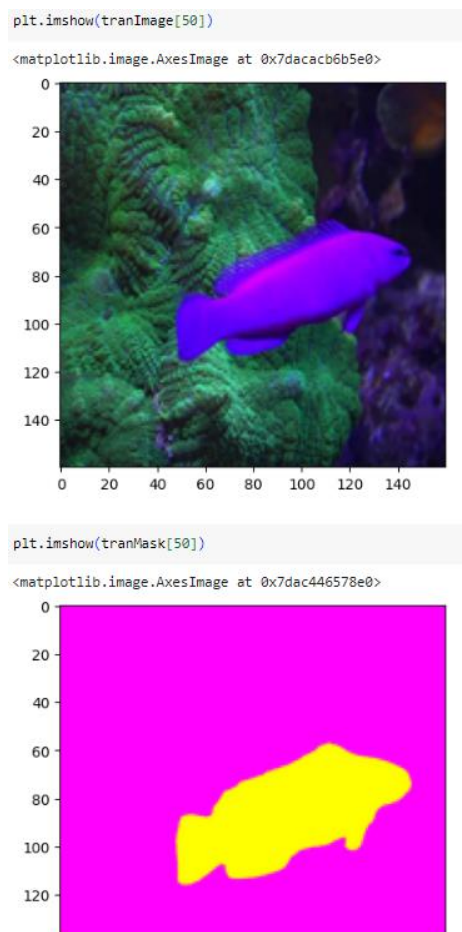
پوشه TEST: شامل ۱۱۰ نمونه جفتی برای ارزیابی عملکرد مدل است. مشابه دایرکتوری Train_val است و احتمالاً شامل زیردایرکتوری‌های تصاویر و ماسک‌ها می‌باشد. داخل دایرکتوری TEST:

images: تصاویر آزمون RGB برای ارزیابی عملکرد مدل.

masks: برچسب‌های حقیقی متناظر با تصاویر آزمون.

ممکن است ماسک‌های RGB ترکیبی برای ارزیابی ارائه شوند.

- پس از بارگیری دیتاست از صحت بارگیری خود و درست خواندن ماسک مرتبط با هر تصویر اطمینان حاصل فرمائید. برای مثال به تصویر زیر توجه فرمائید.



شکل ۱- اطمینان از صحت استخراج دیتا و ماسک مربوط به آن

- از داده‌های تست اختصاص داده شده صرفاً برای تست نهایی مدل در انتهای آموزش شبکه استفاده کنید و داده‌های ارزیابی را از داده‌های آموزشی جدا نمایید. برای این کار ده درصد داده‌ها را جدا نمایید. (لازم است داده‌های آموزشی و ارزیابی تا حد ممکن به صورت متوازن انتخاب بشوند، راجع به علت این تصمیم مطالعه فرموده و مزایای آن را بگوئید)

- پیش از استفاده از مجموعه داده در آموزش و ارزیابی شبکه، حتما داده‌ها را نرمال کنید و روش بکارگرفته شده و علت انتخاب آن جهت نرمالسازی را شرح دهید.
- در صورت نیاز می‌توانید از تقویت داده (data augmentation) مناسب جهت افزایش دقت شبکه و جلوگیری از بیش‌برازش (Overfitting) استفاده نمائید. (لازم است چندین بار انواع مختلف تقویت داده‌ها را پیاده کرده تا بهترین را بنا بر تاثیری که روی شبکه می‌گذارد برگزینید و سپس راجب تقویت داده منتخب توضیح دهید)

۲-۴. شبکه مورد استفاده

(۱۰ نمره)

- سازوکار و معماری دو شبکه UNet و TA-Unet را از روی مقاله ارسال شده مطالعه کنید و هر کدام را توضیح دهید.

۳-۴. آموزش شبکه

(۲۰ نمره)

- شبکه‌های ذکر شده را با رعایت تمامی نکات پیاده سازی نمائید و با مجموعه داده SUIM آموزش دهید.
- ابرپارامترهای (hyperparameters) متناسب با هر مدل را استفاده کنید.

۴-۴. ارزیابی و تحلیل نتایج

(۴۰ نمره)

- راجب مقیاس Mean Intersection over Union (mIoU) مطالعه فرموده و توضیح دهید.
- در این قسمت سوال به ازای هر آموزش انجام شده نمودار mIOU و تابع هزینه داده‌های آموزشی و ارزیابی در حین آموزش شبکه را ترسیم کنید.
- مقدار mIOU شبکه بر روی داده‌های تست برای هریک از شبکه‌های آموزش داده شده را گزارش نمائید.
- در انتها لازم است تا بهترین شبکه را برای بخش‌بندی (Segmentation) انتخاب نموده و علت انتخاب خود را ذکر کنید.

