

توضیح مدل U-net:

1. Encoder (پایین‌رو یا Contracting Path)

- ورودی: تصویر با اندازه 224×224 و 3 کانال رنگی
 - عملیات: ترکیبی از 2 فیلتر 3×3 Conv2D و MaxPooling2D با $\text{stride}=2$ برای استخراج ویژگی‌ها که downsampling انجام داده
 - خروجی: نگاشت‌های ویژگی (feature maps) فشرده‌شده
- ☐ بلوک‌های c1 تا c5 لایه‌های Convolution هستند
 - ☐ p1 تا p5 عملیات Pooling هستند
-

2. Bottleneck

- دو لایه کانولوشن با 1024 فیلتر
 - عمیق‌ترین بخش مدل برای درک ویژگی‌های سطح بالا
-

3. Decoder (بالارو یا Expanding Path)

- از Conv2DTranspose برای بزرگ‌نمایی تصویر استفاده شده (UpSampling)

- در هر مرحله، ویژگی‌ها از مسیر encoder نیز به صورت skip connection افزوده می‌شوند (با concatenate)

- هدف: بازیابی دقت فضایی (spatial resolution)

- بلوک‌های u8 تا u12 لایه‌های upsampling
- بلوک‌های c8 تا c12 عملیات بازسازی (reconstruction) هستند

Output .4

- یک لایه Conv2D نهایی با فیلتر 1×1 و activation برابر softmax برای تقسیم‌بندی چندکلاسه (21 کلاس)

توضیح jaccard_coef:

یک امتیاز شباهت است. به طور کلی بین 0 تا 1 است و هر چه به یک نزدیک تر باشد دو مجموعه مورد بررسی شباهت بیشتری دارند. Jaccard_coef و IoU فرمول‌های مشابهی دارند.

نتایج:

متأسفانه به دلیل زمان بر بودن بسیار فرایند یادگیری نتوانستیم که ران را تمام کنیم. همیشه با ارور runtime برخوردیم. ولی طبیعتاً باید قسمتی که بدون فریز لایه انکودر هست، دقت بیشتری داشته باشد.