

توضیح مدل U-net :

(Contracting Path Encoder .1

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

Bottleneck .2

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

(Expanding Path Decoder .3

- 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 نزدیک‌تر باشد دو مجموعه مورد بررسی مشابهت بیشتری دارند. **Jaccard_coef** و **IoU** فرمول‌های مشابهی دارند.

:نتایج

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