

NOTE

(1) pooling Layer { size 1/2
| channel 不变

下采样

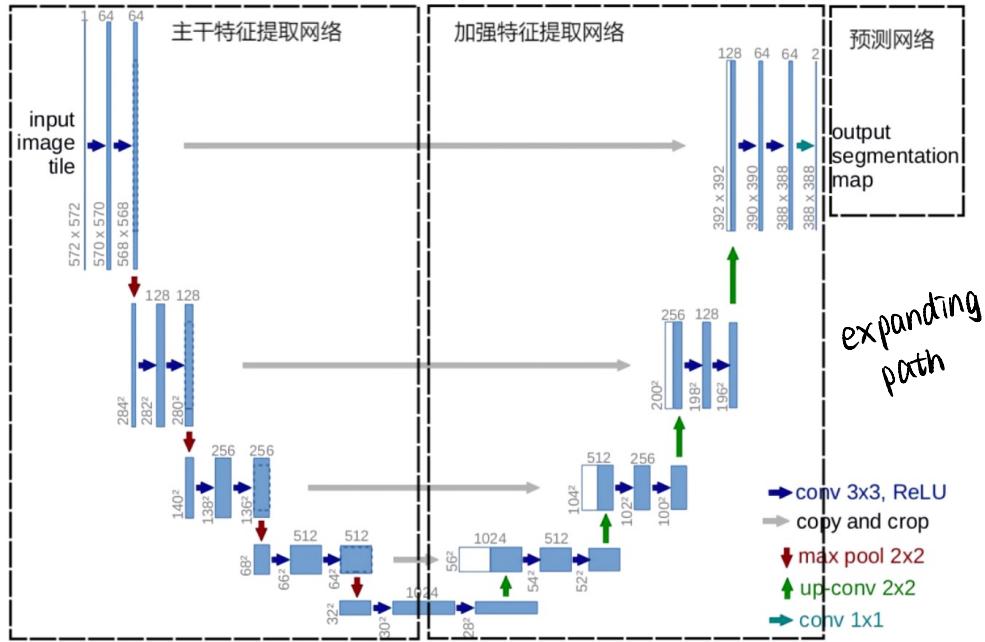
240906 53

(2) Transconv { size X2
| channel 1/2

上采样

part 01 U-Net 网络模型结构

contracting path



(1) U-Net { contracting path (收缩路径)
expanding path (扩展路径)

(2) 训练过程

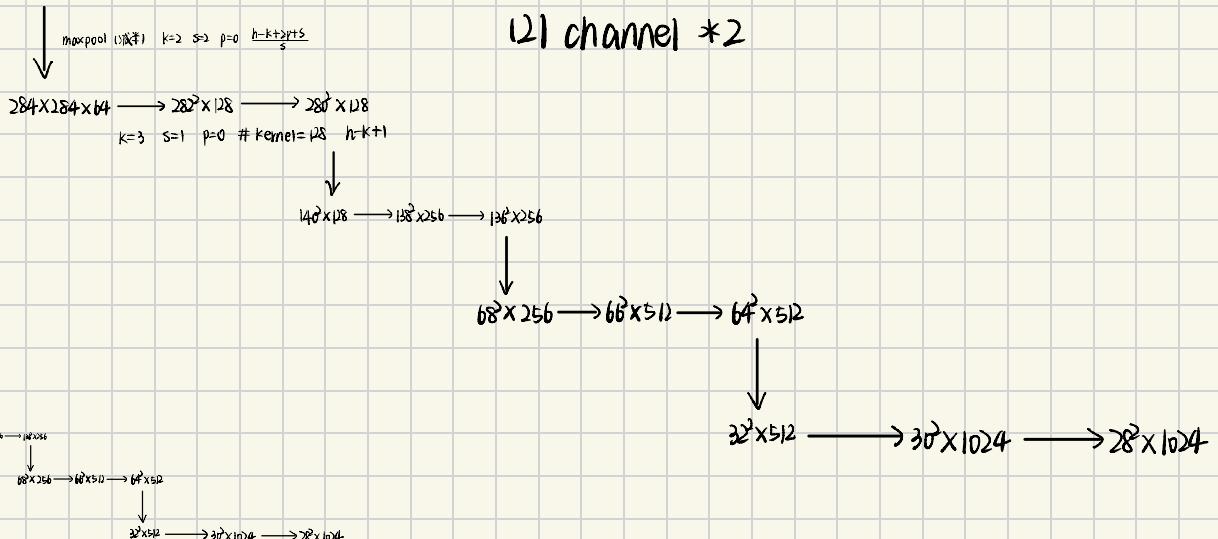
● Contracting path = 4组 $2 \uparrow 3 \times 3 \text{ conv} + 1 \uparrow \text{pooling layer}$ 组成
pooling Layer 下采样

$572 \times 572 \times 1$ $\longrightarrow 570 \times 570 \times 64$ $\longrightarrow 568 \times 568 \times 64$
Input image $k=3 \quad s=1 \quad p=0 \text{ (no padding)} \quad \# \text{ kernel} = 64 \quad h-k+1$

maxpool (减半) $k=2 \quad s=2 \quad p=0 \quad \frac{h-k+2p+s}{s}$
 \downarrow
 $284 \times 284 \times 64$

570x570x1 → 570x570x64 → 568x568x64
 Input image k=3 s=1 p=0 (no padding) # kernel=64 h=k+1

(1) size //2 max_pool



expanding path = 4组

中N裁剪 & 拼接 + 1个上采样层 (transconv) + 2个 3x3 conv

构成

note: copy & crop 中N裁剪 + 拼接

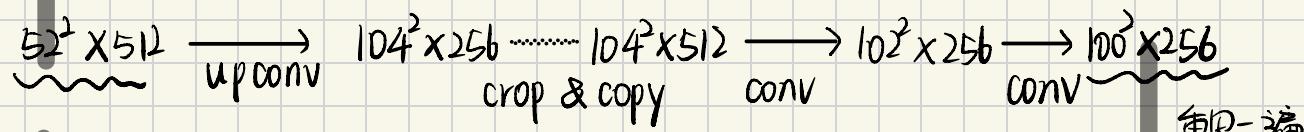
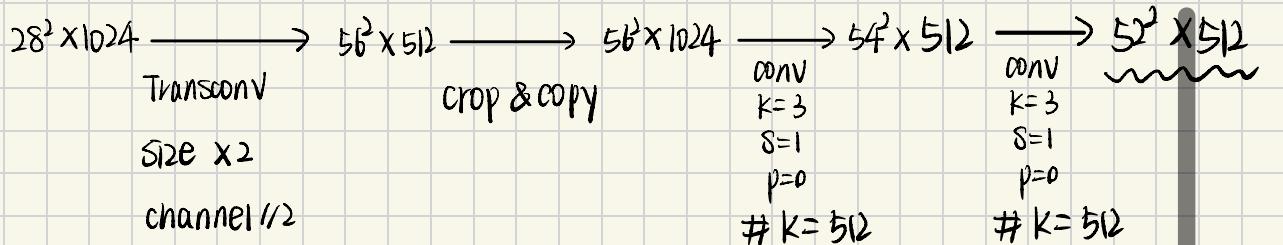
先裁剪

64x64x512 → crop (中N裁剪) → 56x56x512

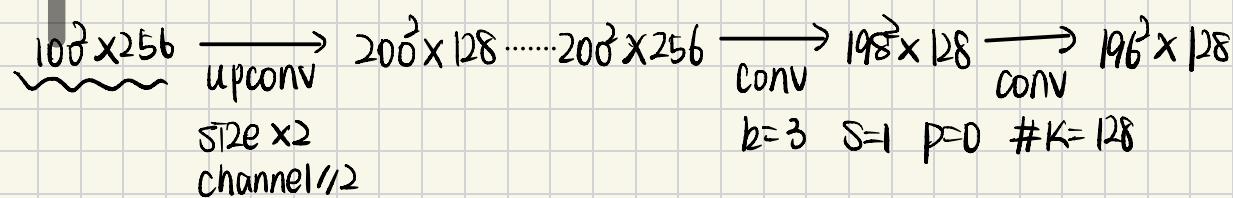
裁剪的是 contracting path (收缩路径) 中的一个 feature map

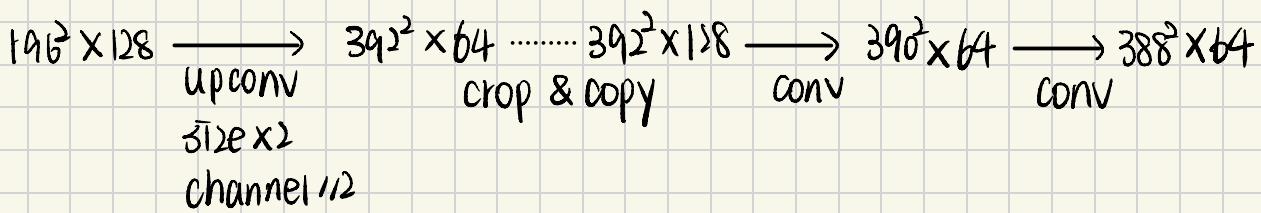
再在 expanding path (↑ 展路径) 中进行拼接

扩展路径的起点



重B-遍





Finally

$$\begin{array}{ccc}
 388^2 \times 64 & \xrightarrow{\text{conv}} & 388^2 \times 2 \\
 k=1 & & \Delta \\
 s=1 \quad p=0 & & \\
 \# k=2 = \text{num_cls} & &
 \end{array}$$

U-Net网络由两部分构成，contracting path（收缩路径）和expanding path（扩展路径）。

U-Net网络训练过程：

1. contracting path（收缩路径）：由4组{两个3x3卷积层+一个池化层（下采样）}构成。

输入特征图 $(572 \times 572 \times 1)$ \rightarrow 长、宽、通道数 $(570 \times 570 \times 64)$ \rightarrow $(568 \times 568 \times 64) \rightarrow$ max_pooling(池化)(减半) \rightarrow $(284 \times 284 \times 64)$ ，两个卷积层+一个池化层.....最后到特征图 $(32 \times 32 \times 512)$ 。

2. 中间又经过两个3x3卷积层：特征图 $(32 \times 32 \times 512) \rightarrow (30 \times 30 \times 1024) \rightarrow (28 \times 28 \times 1024)$

3. expanding path（扩展路径）：由4组{中心裁剪和拼接+一个上采样层（转置卷积）+两个3x3卷积层}构成。

注意：copy and cope 中心裁剪和拼接，先进行裁剪 $(64 \times 64 \times 512) \rightarrow (56 \times 56 \times 512)$ 。这里裁剪的是contracting path（收缩路径）中的一个特征图。再在expanding path（扩展路径）中进行拼接。

特征图 $(28 \times 28 \times 1024) \rightarrow$ up-conv(上采样, 转置卷积) $\rightarrow (56 \times 56 \times 512) \rightarrow$ cope(拼接, 上面中心裁剪得到的特征图) $\rightarrow (56 \times 56 \times 1024) \rightarrow$ conv(3x3卷积) $\rightarrow (54 \times 54 \times 512) \rightarrow$ conv(3x3卷积) $\rightarrow (52 \times 52 \times 512)$ ，一次中心裁剪+一个上采样层（转置卷积）+两个卷积层.....最后得到特征图 $(388 \times 388 \times 64)$ 。

4. 最后进行一次1x1卷积：特征图 $(388 \times 388 \times 64) \rightarrow$ 特征图 $(388 \times 388 \times 2)$ 。最后输出一个 $388 \times 388 \times 2$ 的分割图。

U-Net网络模型改进：在步骤2和步骤3中的卷积层改为大小为3x3，填充为1的卷积层，这样expanding path（扩展路径）中的特征图经过上采样后的大小与contracting path（收缩路径）中对应的特征图大小一致，可以省去中心裁剪这一步直接拼接。



https://blog.csdn.net/qq_45981086/article/details/130439136 语义分割学习笔记（五）U-Net网络 - U-Net网络模型 - CSDN博客

U-Net 2015 生物医学影像 (VGG-12层)

Funny 6

(1) encoder - decoder

Get B

(2) |
 | contracting path
 | expansive path

(3) note: 输入 & 输出不同

(4) $572 \times 572 \times 1$ 单通道输入图像为例

no padding $S=1$ $k=3$

$$572 - 3 + 1 = 570$$

没有 Batch Normalization

(5) max pooling 2×2 $S=2$

pool 操作前后图像大小?

$$H \rightarrow 2p - k + 1$$

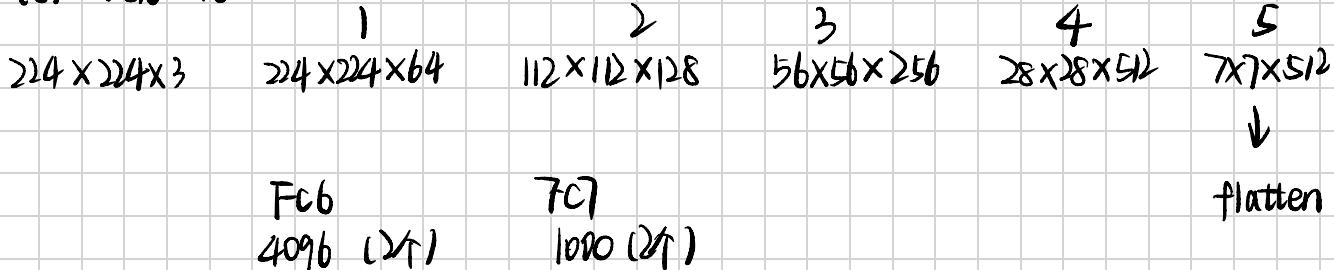
$$\frac{H + 2p - k + s}{s} \quad S=2 \quad k=2$$

∴ 图像尺寸减半

pool channel 不会变的

conv channel 翻倍

(6) VGG-16



(7) up-conv

高 & 宽放大两倍 Trans conv

Tn $28 \times 28 \times 1024$

deal: Trans conv

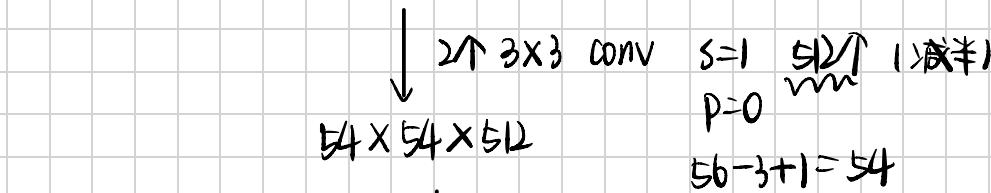
out: $56 \times 56 \times 512$

↓
图片尺寸大，channel 减半

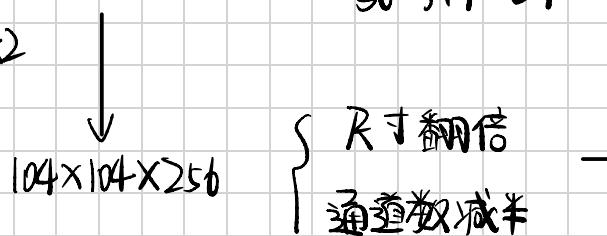
18) copy and crop



17) & 18) $\longrightarrow 56 \times 56 \times 1024$



19) up-conv 2x2



110) skip connection

$\left. \begin{array}{l} \text{copy \& crop} \\ \text{中川裁剪} \end{array} \right\}$

concat
 104×512

$136 \times 136 \times 256$

$\longrightarrow 104 \times 104 \times 256$

$P=0 \quad S=1$
 $3 \times 3 \text{ conv}$

102×256

$P=0 \quad S=1$
 $3 \times 3 \text{ conv}$

100×256

\downarrow
 200×128

111) up-sample

transconv

$\left. \begin{array}{l} \text{size } \times 2 \\ \text{channel } 1/2 \end{array} \right\}$

112) skip-connection

280×128

$\xrightarrow{\text{中川裁剪}} 200 \times 128$

concat
 200×256

同一个TR用
 \uparrow
 $\left. \begin{array}{l} \text{size 缩减 } \times 1 \\ \text{通道减半 } 1/2 \end{array} \right\}$

用同一个卷积核

\Leftrightarrow 连续用 2个 3×3 conv

\downarrow
 198×128

\downarrow
 198×128
 $2 \uparrow 3 \times 3 \text{ conv}$

(13) finally $388 \times 388 \times 64$ $\xrightarrow[1 \times 1 \text{ conv}]{\text{kernel} = \text{num_cls}}$ $388 \times 388 \times 2$
 1x1 conv 特点 $\left\{ \begin{array}{l} \text{size 不变} \\ \text{channel 变少} \end{array} \right.$ $\left\{ \begin{array}{l} \text{在语义中有2个类别} \\ \text{前景 \& 背景} \\ (\# \text{kernel} - \text{size} = 2) \end{array} \right.$

Note: 我们 input 572×572

分割图	388×388	不变 H & W
现在我们的修改	$\left\{ \begin{array}{l} \text{padding} \\ \text{BN layer} \end{array} \right.$	$\left\{ \begin{array}{l} \text{same padding} \\ \text{CONV - BN - ReLU} \end{array} \right.$
	$\left\{ \begin{array}{l} \text{size 不变} \\ \text{channel 变少} \end{array} \right.$	<u>same kernel 连接用 2xR</u>

eg: $64^2 \times 512$
 \downarrow max-pool 2×2 $\left\{ \begin{array}{l} \text{size } 1/2 \\ \text{channel 不变} \end{array} \right.$

看那个底

$32^2 \times 512$
 \downarrow conv 3×3 BN ReLU $\left\{ \begin{array}{l} \text{same padding} \\ s=1 \end{array} \right.$
 $32^2 \times 1024$
 \downarrow 2×2 串联 $1024 @ 3 \times 3$ conv
 $16^2 \times 1024$
 \downarrow up-conv 2×2 $\left\{ \begin{array}{l} \text{size } \times 2 \\ \text{channel } 1/2 \end{array} \right.$
 $64^2 \times 512$
 \downarrow skip connection
 $64^2 \times 1024$ 不需要中 101 裁剪了
 直接 concat 拼接

如何让 output size = input size

* 目前的主流方法

What is U-Net? Done?

把整个网络的输入 & 输出 deal 得清楚就好。

U-Net training process

(1) U-Net { contracting path 收缩路径
expanding path 扩展路径

(2) contracting path

- ① Input Image size $572 \times 572 \times 1$ 2个连续的 3×3 卷积 $s=1$ $p=0$ # $k=64$
size 分别为 $570 \times 570 \times 64$, $568^2 \times 64$
- ② 接下来, 对输出的 feature map 进行 maxpool (下采样) $k=2$ $s=2$ $p=0$ # $k=64$
输出 $284^2 \times 64$
- ③ 进行连续的2个 3×3 conv $s=1$ $p=0$ # $k=128$
分别输出 $282^2 \times 128$ $280^2 \times 128$
- ④ maxpool (下采样) $k=2$ $s=2$ $p=0$ # $k=128$
输出 $140^2 \times 128$
- ⑤ 进行连续的2个 3×3 conv $s=1$ $p=0$ # $k=256$
输出分别为 $138^2 \times 256$, $136^2 \times 256$
- ⑥ maxpool (下采样) $k=2$ $s=2$ $p=0$ # $k=256$
输出 $68^2 \times 256$
- ⑦ 连续的2个 3×3 conv $s=1$ $p=0$ # $k=512$
输出分别为 $66^2 \times 512$, $64^2 \times 512$