

Unet神經網絡為什麼會在醫學圖像分割表現好？

計算機視覺life 2022-03-29 11:02

點擊上方“[計算機視覺life](#)”，選擇“星標”

快速獲得最新乾貨

本文轉載自小白學視覺，作者小白。文章僅用於學術分享。

問題：Unet神經網絡為什麼會在醫學圖像分割表現好？

<https://www.zhihu.com/question/269914775>

Unet網絡在被提出後，就大範圍地用於醫學圖像的分割。其能在醫學圖像展現優秀的性能和它本身網絡結構存在怎樣的一種聯繫？

專業回答

作者：王沈

<https://www.zhihu.com/question/269914775/answer/586501606>

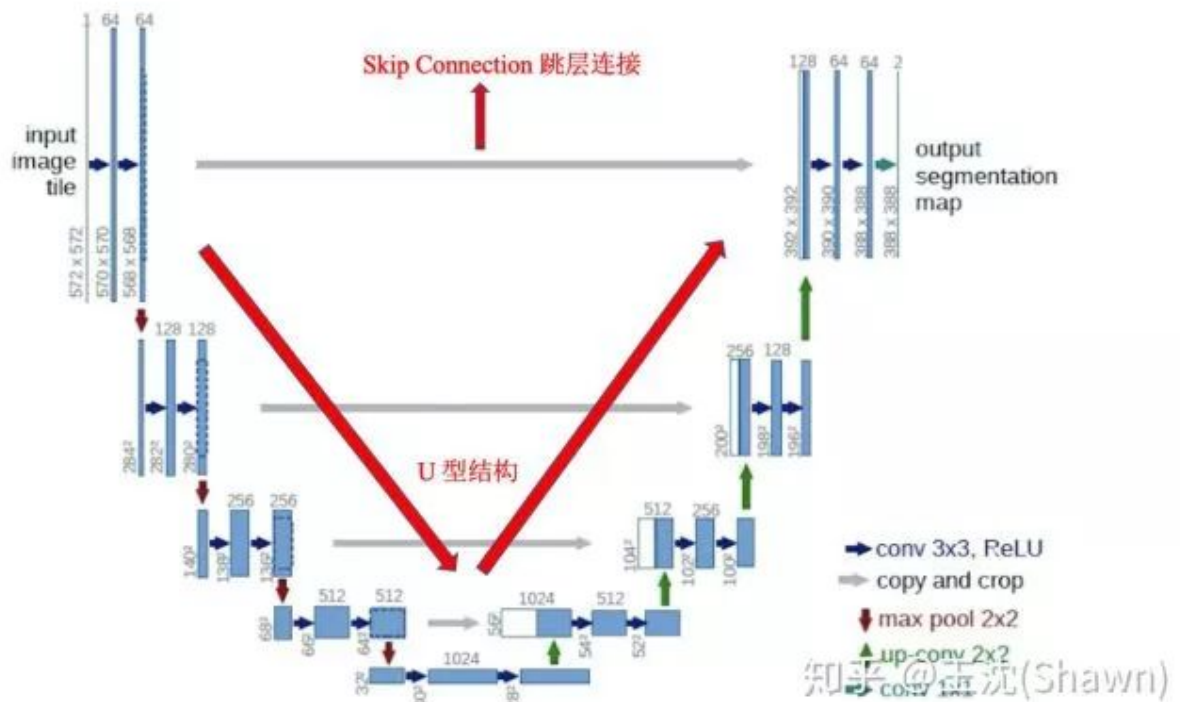
這個問題在面試醫療影像算法崗位的時候，偶爾會提到，我這裡提供一些個人的思考。問題中有兩個關鍵詞，【UNet】和【醫療影像】，接下來我們一一分析這兩個關鍵詞。

首先我們說說【UNet】。

UNet最早發表在2015的MICCAI上，短短3年，引用量目前已經達到了4070，足以見得其影響力。而後成為大多做醫療影像語義分割任務的baseline，也啟發了大量研究者去思考U型語義分割網絡。而如今在自然影像理解方面，也有越來越多的語義分割和目標檢測SOTA模型開始關注和使用U型結構，比如語義分割Discriminative Feature Network(DFN)(CVPR2018)，目標檢測Feature Pyramid Networks for Object Detection(FPN)(CVPR 2017)等。

我們言歸正傳，UNet只是一個網絡結構的代號而已，我們究其細節，到底UNet是由哪些組件構成的呢？

UNet的結構，我認為有兩個最大的特點，U型結構和skip-connection（如下圖）。



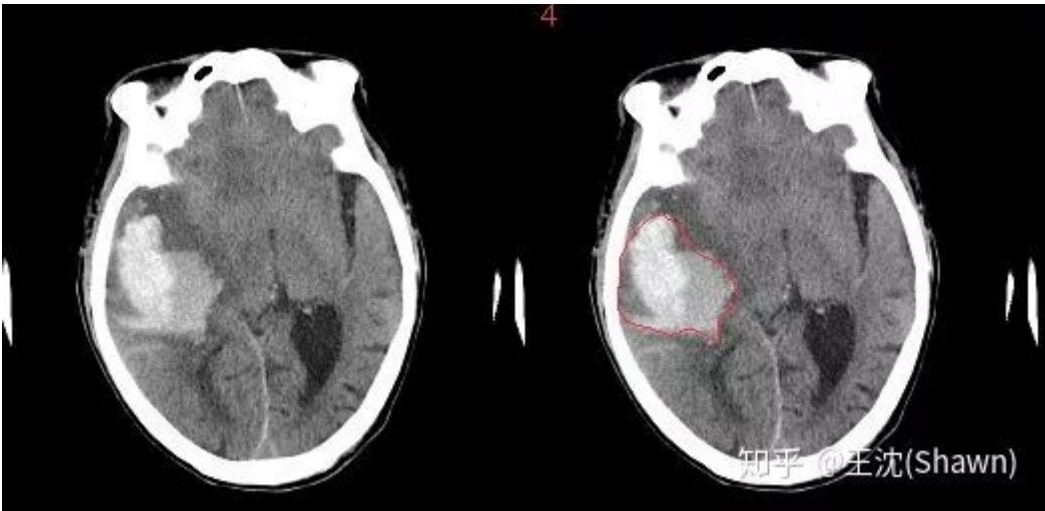
UNet的encoder下採樣4次，一共下採樣16倍，對稱地，其decoder也相應上採樣4次，將encoder得到的高級語義特徵圖恢復到原圖片的分辨率。

相比於FCN和Deeplab等，UNet共進行了4次上採樣，並在同一個stage使用了skip connection，而不是直接在高級語義特徵上進行監督和loss反傳，這樣就保證了最後恢復出來的特徵圖融合了更多的low-level的feature，也使得不同scale的feature得到了融合，從而可以進行多尺度預測和DeepSupervision。4次上採樣也使得分割圖恢復邊緣等信息更加精細。

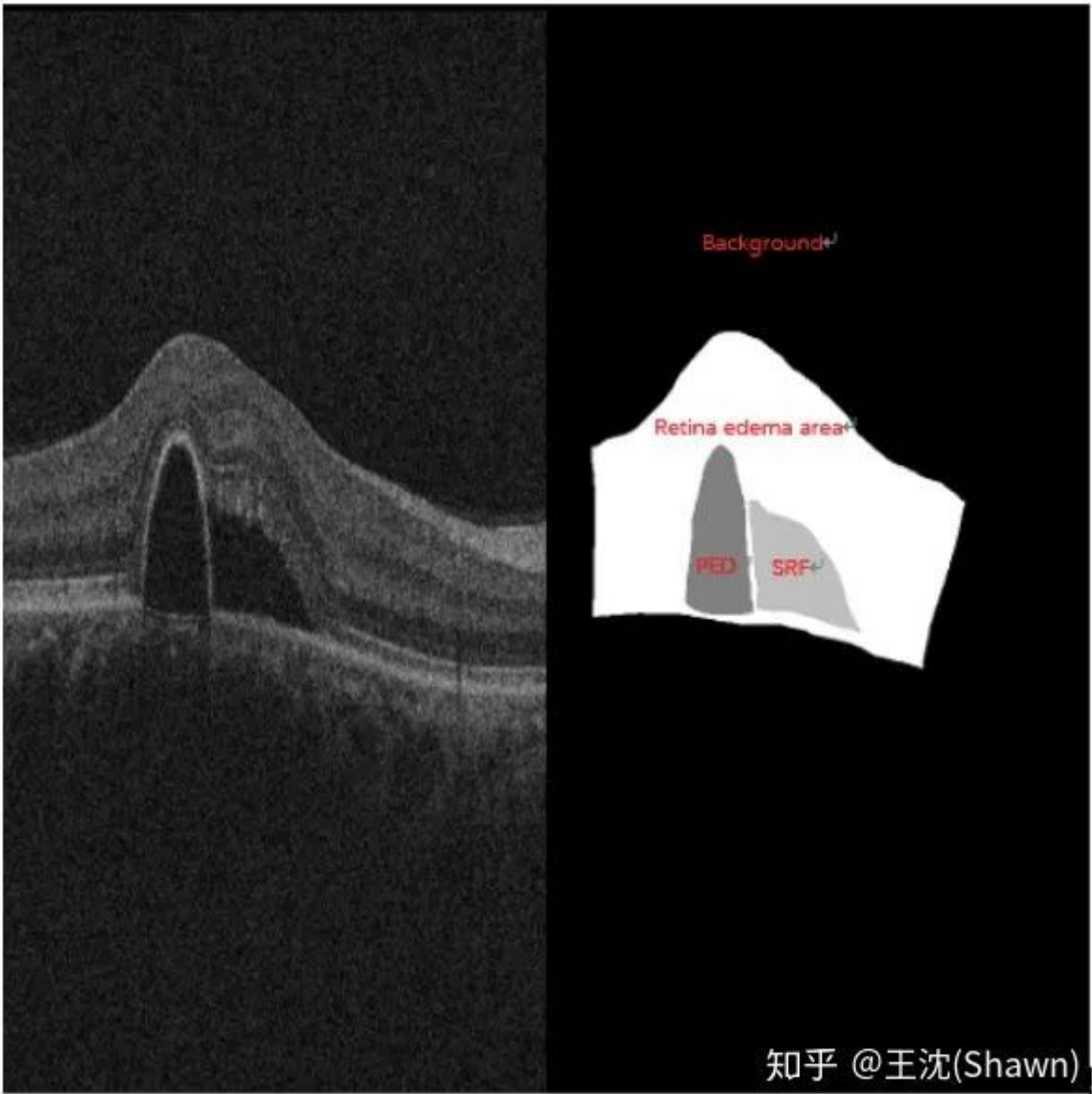
其次我們聊聊【醫療影像】，醫療影像有什麼樣的特點呢（尤其是相對於自然影像而言）？

1. 圖像語義較為簡單、結構較為固定。我們做腦的，就用腦CT和腦MRI，做胸片的只用胸片CT，做眼底的只用眼底OCT，都是一個固定的器官的成像，而不是全身的。由於器官本身結構固定和語義信息沒有特別豐富，所以高級語義信息和低級特徵都顯得很重要(UNet的skip connection和U型結構就派上了用場)。舉兩個例子直觀感受下。

A. 腦出血. 在CT影像上，高密度的區域就大概率是一塊出血，如下圖紅色框區域。



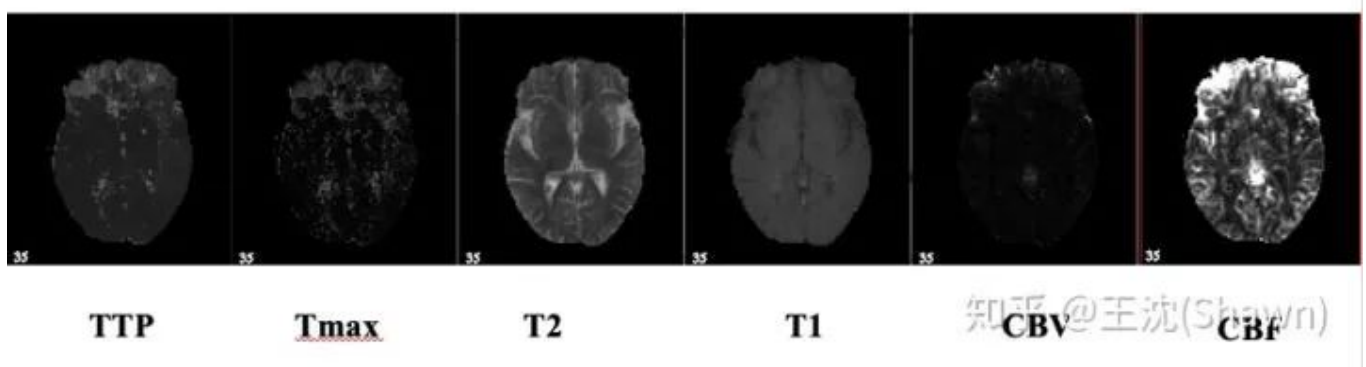
B.眼底水腫。左圖原圖，右圖標註(不同灰度值代表不同的水腫病變區域)。在OCT上，凸起或者凹陷的區域就大概率是一個水腫病變的區域。



2.數據量少。醫學影像的數據獲取相對難一些，很多比賽只提供不到100例數據。所以我們設計的模型不宜多大，參數過多，很容易導致過擬合。

原始UNet的參數量在28M左右(上採樣帶轉置卷積的UNet參數量在31M左右)，而如果把channel數成倍縮小，模型可以更小。縮小兩倍後，UNet參數量在7.75M。縮小四倍，可以把模型參數量縮小至2M以內，非常輕量。個人嘗試過使用Deeplab v3+和DRN等自然圖像語義分割的SOTA網絡在自己的項目上，發現效果和UNet差不多，但是參數量會大很多。

3.多模態。相比自然影像，醫療影像比較有趣和不同的一點是，醫療影像是具有多種模態的。以ISLES腦梗競賽為例，其官方提供了CBF,MTT,CBV,TMAX,CTP等多種模態的數據。



這就需要我們更好的設計網絡去提取不同模態的特徵feature。這裡提供兩篇論文供大家參考。

Joint Sequence Learning and Cross-Modality Convolution for 3D Biomedical Segmentation(CVPR 2017) ,

Dense Multi-path U-Net for Ischemic Stroke Lesion Segmentation in Multiple Image Modalities.

4.可解釋性重要。由於醫療影像最終是輔助醫生的臨床診斷，所以網絡告訴醫生一個3D的CT有沒有病是遠遠不夠的，醫生還要進一步的想知道，病灶在哪一層，在哪一層的哪個位置，分割出來了嗎，能求體積嘛？同時對於網絡給出的分類和分割等結果，醫生還想知道為什麼，所以一些神經網絡可解釋性的trick就有用處了，比較常用的就是畫activation map。看網絡的哪些區域被激活了，如下圖。

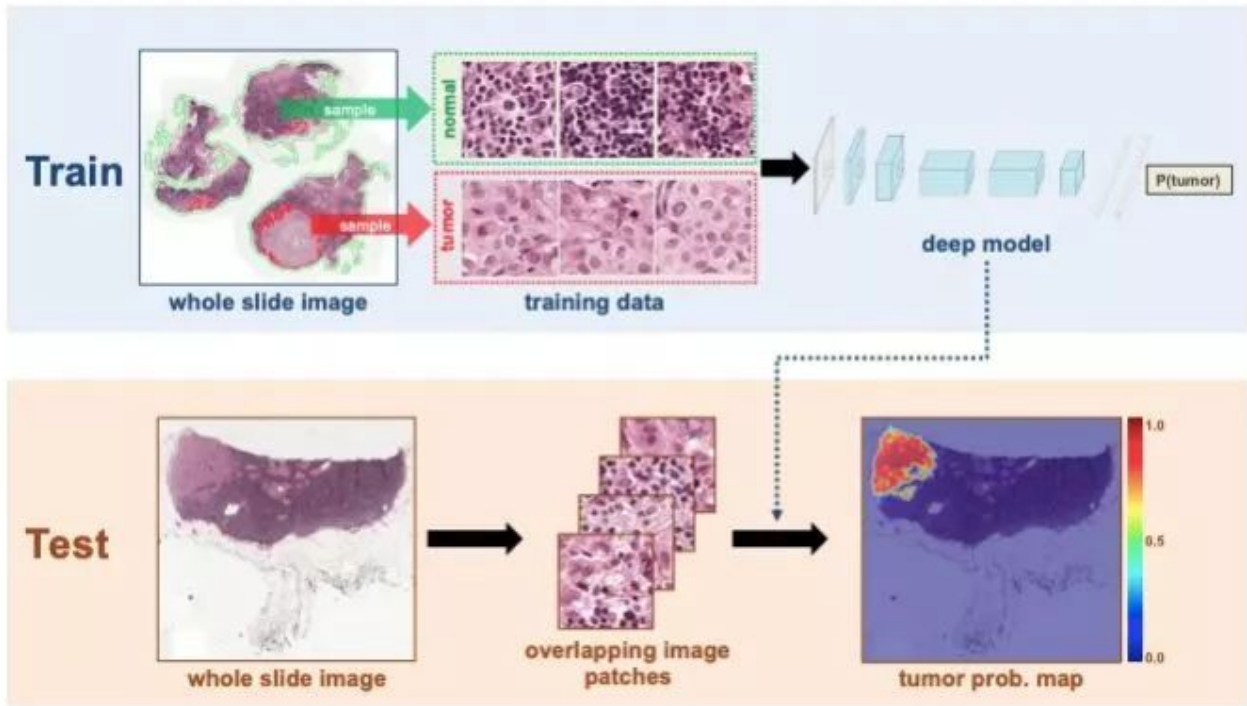


Figure 2: The framework of cancer metastases detection. 知乎 @王沈(Shawn)

這裡推薦兩篇工作：周博磊老師的 Learning Deep Features for Discriminative Localization(CVPR2016)和其實驗室同學的Deep Learning for Identifying Metastatic Breast Cancer（上圖的出處）

（周博磊知乎主頁：<https://www.zhihu.com/people/zhou-bo-lei/activities>）

BTW:沒有偏題的意思，只是覺得醫療影像的特點和本問題息息相關，就一起總結了。

最後提一個問題，引發關注醫療影像的同學們思考和討論？

前面有提到，UNet成為大多做醫療影像語義分割任務的baseline，也啟發了大量研究者去思考U型語義分割網絡。那UNet還有什麼不足呢？歡迎大家關注我的項目UNet-family，尋找答案。

（相關鏈接：<https://github.com/ShawnBIT/UNet-family>）

獨家重磅課程官網：cvlife.net

- 1、機器人導航運動規劃： 機器人核心技術運動規劃：讓機器人想去哪就去哪！
- 2、詳解Cartographer：谷歌開源的激光SLAM算法Cartographer為什麼這麼牛X？
- 3、深度學習三維重建 詳解深度學習三維重建網絡：MVSNet、PatchMatchNet、JDACS-MS
- 4、三維視覺基礎 詳解視覺深度估計算法（單/雙目/RGB-D+特徵匹配+極線矯正+代碼實戰）
- 5、視覺SLAM必備基礎 SLAM的第一個實踐，最適合學哪個開源框架？
- 6、VINS:Mono+Fusion SLAM面試官：看你簡歷上寫精通VINS，麻煩現場手推一下預積分！

- 7、VIO進階：ORB-SLAM3（單/雙目/RGBD+魚眼+IMU+多地圖+閉環）獨家70講！即將關閉購買！
- 8、圖像三維重建課程：視覺幾何三維重建教程（第2期）：稠密重建，曲面重建，點雲融合，紋理貼圖
- 9、重磅來襲！基於LiDAR的多傳感器融合SLAM 系列教程：LOAM、LeGO-LOAM、LIO-SAM
- 10、系統全面的相機標定課程：單目/魚眼/雙目/陣列相機標定：原理與實戰



扫描学习SLAM、三维重建

全國最棒的SLAM、三維視覺學習社區↓



技術交流微信群

歡迎加入公眾號讀者群一起和同行交流，目前有SLAM、三維視覺、傳感器、自動駕駛、計算攝影、檢測、分割、識別、**醫學影像**、**GAN**、**算法競賽**等微信群，請添加微信號chichui502 或掃描下方加群，備註：“名字/暱稱+學校/公司+研究方向”。**請按照格式備註，否則不予通過**。添加成功後會根據研究方向邀請進入相關微信群。請勿在群內發送廣告，否則會請出群，謝謝理解~

投稿、合作也歡迎聯繫：simiter@126.com



— 版權聲明 —

本公眾號原創內容版權屬計算機視覺life所有；從公開渠道收集、整理及授權轉載的非原創文字、圖片和音視頻資料，版權屬原作者。如果侵權，請聯繫我們，會及時刪除。

閱讀原文

喜歡此內容的人還喜歡

ICRA 2022 |多物體三維形狀重建和6D位姿和尺寸估計

計算機視覺life