

Training an encoder-decoder with limited data for fetal femur segmentation from echographic images

Addestramento di un encoder-decoder per la segmentazione del femore fetale da immagini ecografiche

Ollari Ischimji Dmitri

9 ottobre 2023

Indice

1	Introduzione	3
1.1	Binary Semantic Segmentation	3
1.2	Fully Convolutional Network	3
1.3	U-Net	3
2	Related Works	4
2.1	Segmentazione ossea	4
2.2	Segmentazione vasi sanguigni	4

Elenco delle figure

Elenco delle tabelle

1 Introduzione

1.1 Binary Semantic Segmentation

La segmentazione semantica è una tecnica di *computer vision* che permette di assegnare ad ogni pixel di un'immagine un'etichetta che ne descrive il contenuto.

Nello specifico in questa tesi si tratta una sottocategoria della segmentazione semantica, ovvero la **Binary Semantic Segmentation** (segmentazione semantica binaria), questa tecnica di *computer vision* permette di assegnare ad ogni pixel di un'immagine un'etichetta che ne descrive il contenuto, ma a differenza della segmentazione semantica classica, che permette di assegnare ad ogni pixel una delle N possibili etichette, la segmentazione semantica binaria permette di assegnare ad ogni pixel una delle due etichette possibili: **oggetto** o **sfondo**.

La segmentazione è una tipologia di problema molto ricorrente in ambito medico, in quanto permette di automatizzare alcune procedure che altrimenti sarebbero eseguite manualmente, riducendo i tempi di esecuzione e i costi, permettendo di ottenere risultati più precisi e accurati limitando lo sforzo umano.

1.2 Fully Convolutional Network

Le **Fully Convolutional Network** (FCN) [4] sono una tipologia di reti neurali convoluzionali (CNN) che permettono di effettuare segmentazioni semantiche, in quanto sono in grado di gestire input di qualsiasi dimensione e di produrre mappe di segmentazione più precise grazie alla loro capacità di apprendere contesti spaziali.

Le motivazioni riguardanti l'ampio utilizzo nel settore della *computer vision* sono legate all'assenza di strati completamente connessi (lineari) che vincolano l'input alla medesima grandezza per ogni singola immagine, permettendo di fornire in input l'intera immagine e non frammenti della stessa così da aumentare l'apprendimento spaziale della rete.

Questa maggior flessibilità comporta un'addestramento libero da limitazioni sull'input comportando una maggiore tolleranza agli errori e al rumore rendendo questa tipologia di reti particolarmente adatte a contesti poveri di dati.

1.3 U-Net

L'architettura **U-net** [5] è una particolare implementazione di FCN che permette di effettuare segmentazioni semantiche, in quanto è in grado di gestire input di qualsiasi dimensione e di produrre mappe di segmentazione più precise grazie alla sua capacità di apprendere contesti spaziali.

2 Related Works

2.1 Segmentazione ossea

Un progetto degno di nota è **Towards whole-body CT Bone Segmentation** [3] poichè si propone di risolvere il problema della segmentazione ossea di umani e la raggiunge con ottimi risultati con accuratezza del $96\% \pm 2\%$ mediante la metrica di **Dice Score** e $94\% \pm 2\%$ con la metrica di **Intersection over Union**.

Il progetto si basa su una rete neurale convoluzionale che utilizza l'architettura U-Net [5] e raggiunge ottimi risultati con circa 4000 immagini e 60 epoche di training lasciando l'architettura della rete invariata.

2.2 Segmentazione vasi sanguigni

L'articolo **Accurate Retinal Vessel Segmentation via Octave Convolution Neural Network** [2] propone un metodo di segmentazione automatico per la segmentazione dei vasi sanguigni retinici.

L'implementazione del metodo è basata su una rete neurale convoluzionale che utilizza l'architettura Octave UNet, modello che segue l'architettura di U-Net [5] ma utilizza l'operazione di convoluzione octave e delle convoluzioni octave trasposte.

Le convoluzioni octave sono state introdotte in **Drop an Octave: Reducing Spatial Redundancy in Convolutional Neural Networks with Octave Convolution** [1] e sono una variante delle convoluzioni standard che prova a ovviare al problema di sbilanciamento delle *features* all'interno delle mappe di *features*.

Riferimenti bibliografici

- [1] Yunpeng Chen, Haoqi Fan, Bing Xu, Zhicheng Yan, Yannis Kalantidis, Marcus Rohrbach, Shuicheng Yan, and Jiashi Feng. Drop an octave: Reducing spatial redundancy in convolutional neural networks with octave convolution, 2019.
- [2] Zhun Fan, Jiajie Mo, Benzhang Qiu, Wenji Li, Guijie Zhu, Chong Li, Jianye Hu, Yibiao Rong, and Xinjian Chen. Accurate retinal vessel segmentation via octave convolution neural network, 2020.
- [3] André Klein, Jan Warszawski, Jens Hillengaß, and Klaus Hermann Maier-Hein. Towards whole-body CT bone segmentation. In *Bildverarbeitung für die Medizin 2018*, pages 204–209. Springer Berlin Heidelberg, 2018.
- [4] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation, 2015.
- [5] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.