

# Report Biomedical Vision Project

Valentina Marandici

Giugno 2023

## 1 Introduzione

In questo report verranno riportati i metodi utilizzati per risolvere il problema posto dalla Kits19 Challenge. Lo scopo principale della challenge era sviluppare una rete neurale che fosse in grado di segmentare efficacemente i reni e gli eventuali tumori presenti all'interno di un'immagine attraverso il supervised learning. Nel dataset fornito erano presenti immagini biomediche relative a 210 pazienti, acquisite attraverso il metodo della tomografia computerizzata (CT). Inoltre nel dataset venivano fornite le segmentazioni relative a ogni singola immagine, che consentivano di effettuare il supervised learning. Per affrontare il task della segmentazione multiclasse è stata scelta una rete Unet appositamente modificata per poter segmentare diverse classi.

## 2 Dataset

Per ovviare al problema della limitatezza della capacità computazionale a disposizione, il dataset è stato così ristrutturato: dai 210 pazienti ne sono stati scelti 101 e dal momento che ogni immagine presentava numerose fette, ne sono state selezionate solo quelle centrali. In seguito le immagini e le maschere sono state ridimensionate a una dimensione di (256,256) in modo da velocizzare il processo di training e ridurre ulteriormente il carico computativo. Dopo il processo di resizing il dataset è stato così diviso: training, validation e test. Alla fase di training sono state assegnate 40 immagini, mentre per validation e test abbiamo riservato rispettivamente 20 e 21 immagini. Al fine di migliorare la performance

della rete si è scelto di applicare ad alcune immagini e maschere del training la tecnica della data augmentation per ampliare il nostro dataset di training ed evitare il fenomeno dell'overfitting.

### 3 Unet

In seguito a numerose ricerche si è scelto di adottare una rete di tipo Unet, una rete contraddistinta dall'architettura a U, conferitagli da quattro blocchi encoder e quattro blocchi decoder connessi tra di loro attraverso attraverso gli skip-connection. La scelta della Unet è stata presa tenendo conto del fatto che nello stato dell'arte la Unet è stata un'architettura che ha consentito di ottenere un'elevata accuratezza in diversi task di segmentazione di immagini mediche con dataset limitati. Ai primi layers di encoding della unet sono stati aggiunti due regolarizzatori del kernel: L1L2 e L2. Questa aggiunta è dovuta ad una necessità di regolarizzare i pesi all'interno della loss function; infatti, nelle prime prove del modello le due curve della funzione di loss nelle due fasi di training e validation convergevano troppo rapidamente, cosa che è stata notevolmente migliorata dalla nuova aggiunta.

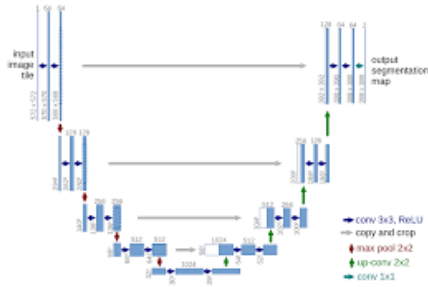


Figure 1: Unet Architecture structure

## 4 Training e Validation

Il dataset di training è composto da 60 immagini originali e 20 immagini a cui è stata applicata la data augmentation. Prima di poter eseguire il modello è stato però necessario modificare le maschere attraverso la tecnica del one-hot encoding. Dal momento che la classe 0, ovvero la classe del background, era di molto prevalente all'interno delle segmentazioni si è deciso di adottare una tecnica per modificare la loss function del modello, ovvero la ricalibrazione dei pesi delle labels, in modo che ai pixel meno prevalenti venisse assegnata una maggiore importanza. Il numero di epoche che meglio si adattava al modello è stato scoperto essere 110.

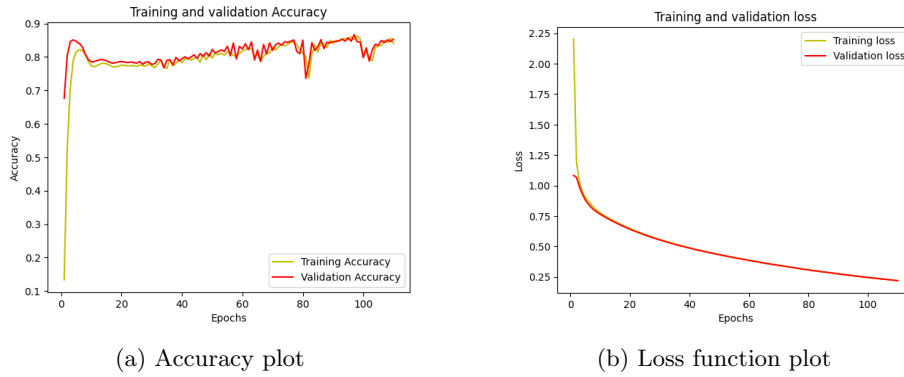


Figure 2: Accuracy and Loss Function Plots

## 5 Conclusione

Dal momento che in seguito al testing il modello si è rivelato altamente impreciso alcuni metodi per migliorarlo potrebbero essere una diversa metodologia nei bilanciare i pesi delle labels o scegliere di utilizzare dei pesi di un modello che è già stato sottoposto al training.