

# Deep learning e segmentazione per la biologia cellulare

Mastrofini Alessandro<sup>1</sup>

<sup>1</sup>alessandro.mastrofini@alumni.uniroma2.eu

Università degli studi di Roma Tor Vergata

Questo lavoro è stato svolto durante il corso *Elaborazione di Immagini*

url: <https://alessandromastrofini.it/deep-learning-segmentazione>

**ABSTRACT** Nell'era del digitale anche in medicina ogni giorno sono prodotte numerose immagini e sebbene la loro classificazione risulta ampiamente studiata la segmentazione e il riconoscimento di oggetti sono meno approfonditi. Con il recente sviluppo dell'intelligenza artificiale e degli algoritmi delle reti neurali convoluzionali è possibile ottenere ottimi risultati. In questo articolo vengono analizzate le differenti reti neurali in Matlab e la possibilità di riconoscere cellule in microscopia ottica. Viene poi introdotta anche la possibilità di segmentare bordo e zona interna estendono l'algoritmo dalle cellule in microscopia al riconoscimento di nei per la dermatoscopia.

**INDEX TERMS** deep-learning, transfer learning, neural network, microscopy, biology, cells, matlab

## INTRODUZIONE

Nella biologia è presente da diversi anni la necessità di automatizzare il riconoscimento cellulare. La possibilità di riconoscere una cellula precisa all'interno di una qualsivoglia immagine in microscopia è un'abilità molto richiesta.

Per decenni tale compito è stato svolto da operatori umani ma con l'aumento esponenziale della potenza di calcolo dei moderni computer è stato possibile introdurre diversi algoritmi di segmentazione. Oggi giorno esistono diverse tecniche per identificare oggetti all'interno delle immagini ma la maggior parte di esse risentono di problemi legati alla qualità, forma, dimensione, convessità e molte altre proprietà geometriche dell'oggetto di interesse. L'ultimo sviluppo vede l'applicazione delle reti neurali e del transfer learning per automatizzare completamente la segmentazione ottenendo prestazioni superiori degli algoritmi tradizionali.

## BACKGROUND

### A. SEGMENTAZIONE

La segmentazione è formalmente definita come "la divisione di un'immagine in un set di regioni non sovrapponibili che unite danno l'intera immagine". Ovvero è un insieme di tecniche per veicolare ed estrarre l'informazione dalle immagini attraverso alcune proprietà, tra cui quelle morfologiche assegnando ad ogni pixel un'etichetta univoca, associata ad un insieme di classi.

Esistono diversi metodi e principi di segmentazione e in tutti si parte dividendo foreground e background. Il fore-

ground è proprio ciò che risulta di interesse e quindi lo si vuole separare dal background. Esistono quattro categorie principali di metodi:

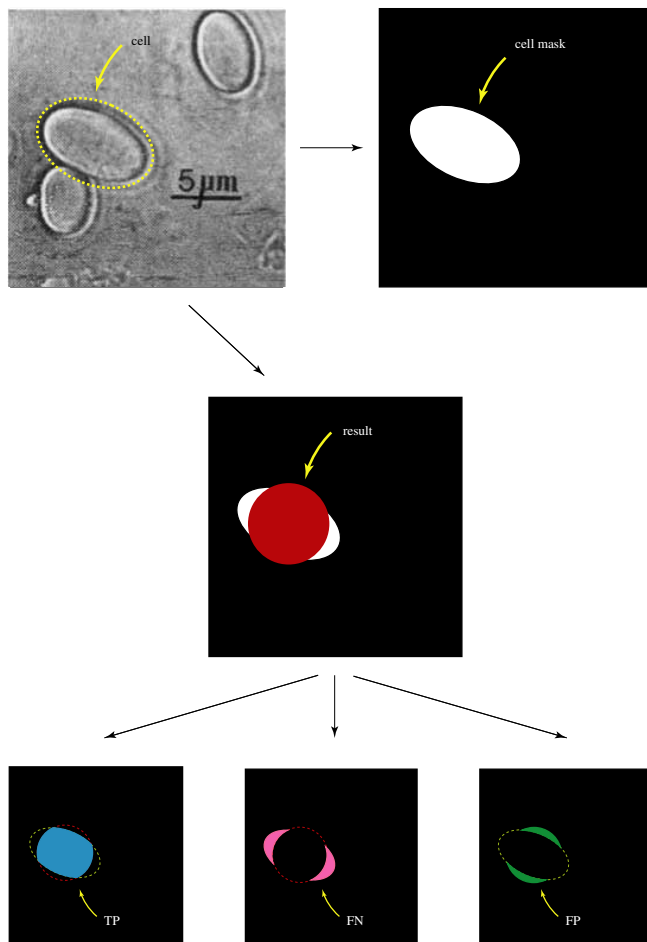
- Pixel-based, basati sulla luminanza dei singoli pixel
- Edge-based, sfruttano bordi e criteri di discontinuità
- Region-based, utilizzano criteri di somiglianza tra regioni vicine
- Model-based, richiedono un modello geometrico dell'oggetto da cercare

A questi si aggiungono i metodi supervisionati che richiedono reti neurali e transfer learning per automatizzare il riconoscimento.

### B. METRICHE DI VALUTAZIONE

Oltre all'immagine di partenza è necessario avere un ground truth (GT) ovvero una maschera di verità che esprime quale pixel dell'immagine è effettivamente un pixel dell'oggetto di interesse. Sulla base di questa maschera e dell'identificazione che segue la segmentazione vengono definite alcune metriche di qualità.

Ci saranno dei pixel falsi positivi (FP) ovvero pixel assegnati all'oggetto che in realtà non ne fanno parte. I veri positivi (TP) invece saranno i pixel dell'oggetto che effettivamente vengono riconosciuti. Infine, ci sono anche i falsi negativi (FN) ovvero i pixel dell'oggetto che vengono persi. Sulla base di queste tre quantità si possono definire due metriche principali ??.



**Figure 1.** Segmentazione ed identificazione di una cellula all'interno di un'immagine in microscopia. Sono segnalati anche un possibile risultato di una segmentazione, il GT e le differenti tipologie di pixel (true positive, false positive, false negative)

a completezza (CM) indica quanto la segmentazione ha effettivamente preso dell'oggetto di partenza mentre la correttezza (CR) tiene conto della sovrasedgmentazione. Sulla base di questi due si introduce il parametro F-measure (FM), eq. (2), che tiene conto sia della sotto segmentazione che della sovra segmentazione. Tale parametro vive nel range [0; 1] ed è tanto migliore quanto più vicino a 1.

$$CM = \frac{TP}{TP + FN} = \frac{TP}{\text{Total area in GT}} \quad (1)$$

$$CR = \frac{TP}{TP + FP} = \frac{TP}{\text{Total area in BW}}$$

$$FM = \frac{2 \cdot CM \cdot CR}{CM + CR} \in [0; 1] \quad (2)$$

### C. DEEP LEARNING

Oggi giorno le reti neurali e l'apprendimento automatico sono molto sviluppate visto anche il recente sviluppo permesso dall'aumento delle potenze di calcolo. Gli ultimi sviluppi che hanno portato all'introduzione sul mercato consumer di computer molto potenti e fotocamere ad alta risoluzione

hanno portato all'utilizzo di reti neurali per elaborare le immagini.

L'apprendimento automatico nasce in contrapposizione alla programmazione tradizionale nell'ottica di mettere la macchina nelle condizioni di estrarre autonomamente alcune features di interesse dai dati. A tale scopo nasce quello che prende il nome di deep learning. Diverse tecniche basate su reti artificiali neurali prevedono l'organizzazione su diversi strati ed ognuno di essi calcola dei valori che a loro volta vengono passati al successivo.

### D. CONVOLUTIONAL NETWORKS

Dati visuali e più generici dati bidimensionali vengono spesso elaborati con reti neurali convoluzionali, CNN. Sono reti composte da uno o più strati convoluzionali con un'architettura feed-forward section -D. Ovvero le diverse connessioni tra le unità non formano cicli ma le informazioni si muovono su una sola direzione rispetto i nodi di ingresso.

Quindi una CNN non è altro che un algoritmo che prende in ingresso un'immagine e fornendo una certa importanza ad alcuni suoi aspetti. Analizzando diversi aspetti di diverse immagini è in grado di distinguere l'una dalle altre. Le diverse reti CNN hanno alcuni punti in comune nella loro architettura.

Partendo dal canale di ingresso questo accetterà immagini di una certa dimensione e con un certo numero di canali. Successivamente l'immagine viene passata ad un primo livello convoluzionale dove avviene un primo filtraggio. Questo porta ad estrarre features di alto livello come ad esempio bordi i bruschi cambi di luminosità.

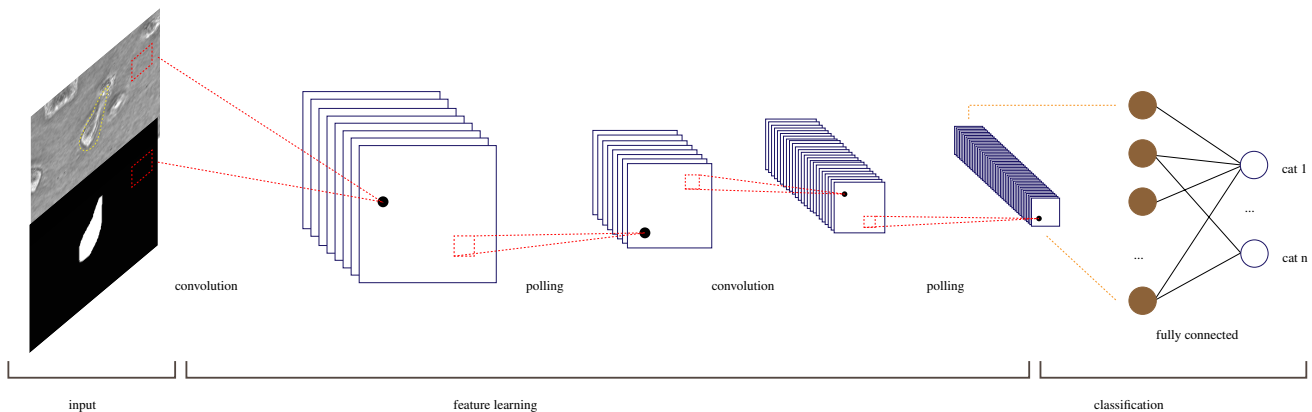
Successivamente ci sono dei layer di pooling dove vengono preservate le features che risultano invarianti e posizione e rotazione andando anche a filtrare parte del rumore. I livelli ReLU (rectified Linear Units) hanno l'obiettivo di annullare ed eliminare i valori non utili dopo che i dati sono passati per un livello convoluzionale. Ci sono poi i layer FC (fully connected) che sono quelli che si occupano della classificazione vera e propria.

### E. TRANSFER LEARNING

È evidente la complessità delle reti neurali e infatti un addestramento accurato richiede un'elevata complessità computazionale. Esiste però una tecnica per adattare un'intelligenza artificiale ad un task diverso da quello per cui è già stata addestrata. L'idea è quella di prendere gli strati inferiori di una rete già addestrata e aggiungere nuovi strati finali adattandoli al problema di interesse. Si utilizzando quindi nuovi set di training e algoritmi di ottimizzazione per adattare la classificazione al nuovo problema.

### ADDESTRAMENTO

A seguire un'analisi su un'implementazione base all'interno di Matlab di alcune reti neurali per la segmentazione di cellule. Verranno valutate anche le prestazioni di addestramento e di segmentazione. All'interno di Matlab è disponibile il Deep Learning Toolbox [1] che fornisce una



**Figure 2.** Schema generico di una rete convoluzionale

piattaforma per la progettazione e l'implementazione di reti neurali profonde con algoritmi e modelli pre addestrati.

#### F. DATASET

Il primo passo per addestrare correttamente una rete neurale è quello di selezionare le immagini, le classi e le maschere di identificazione di ciò che si vuole rappresentare. Ovvero si devono considerare le immagini di riferimento e disporre di una maschera che permette di identificare gli oggetti di interesse al suo interno. In questo caso parliamo di un problema binario, cellula e background, quindi si considerano maschere binarie, 0 sul background e 1 sul background.

A livello computazionale in Matlab si riferisce il tutto ad una variabile di datastore. Ovvero un'unica variabile che al suo interno contiene informazioni sul nome dei file, il loro percorso, ecc senza importarli nel workspace e quindi senza occupare memoria. L'intera collegamento verrà poi processata e caricata solo al momento dell'effettivo addestramento.

#### G. TRAINING

Il primo passo per poter utilizzare efficacemente una rete neurale preaddestrata è quello di addestrarla correttamente. Addestrare una rete consiste nel fornire un numero adeguato di coppie contenenti l'immagine di test e il risultato della segmentazione. Questo permette alla rete di modificare diversi parametri e coefficienti così da poter svolgere i compiti richiesti.

In particolare, in questo caso si vuole fare una segmentazione binaria, separando il foreground dal background. Quindi l'obiettivo è quello di identificare la cellula e distinguere dallo sfondo. Nelle prime analisi viene utilizzato un dataset di riferimento diviso in due grandi insiemi. Una prima parte viene utilizzata per l'addestramento e una seconda parte per una successiva analisi della rete.

#### H. RESNET50

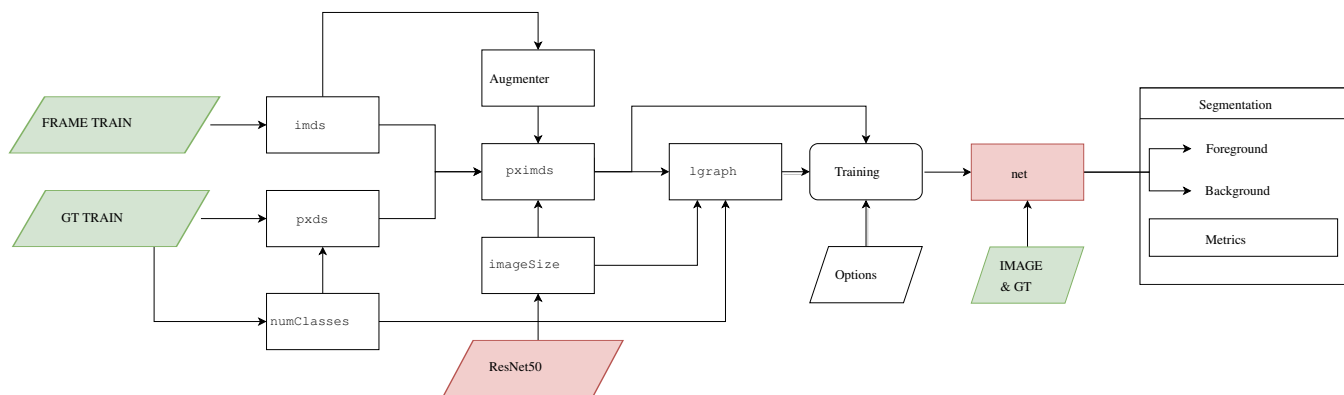
La prima rete che verrà testata è ResNet50, una rete convoluzionale con 50 layers di profondità. La rete pre addestrata è in grado di classificare oltre 1000 categorie di oggetti.

Presenta un totale di 177 layer. Dal layer di input è possibile anche vedere che la rete richiede immagini a 3 canali, rgb, con dimensione 224 x 224. Questo sarà da tenere presente quando viene preparato il dataset di addestramento/test.

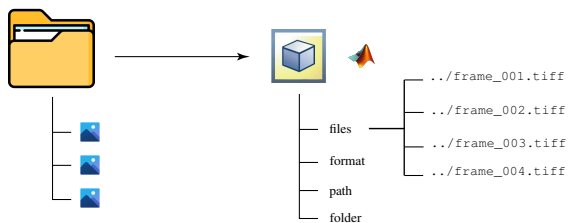
Per estrarre ulteriori informazioni è possibile andare a leggere la struttura della rete. In alternativa è possibile aprire il Deep Network Analyzer ed esplorare i differenti layer nella più evidente complessità della rete.

#### References

- [1] Mathworks. Image processing using deep learning. [Online]. Available: <https://it.mathworks.com/help/deeplearning/image-processing-using-deep-learning.html>



**Figure 3.** Schema della preparazione al training



**Figure 4.** Generico datastore