

POLITECNICO DI TORINO

Facoltà di Ingegneria
Corso di Laurea Magistrale in Ingegneria Biomedica



**Politecnico
di Torino**

Multi-Timepoint Growth Tracking of 3D Cancer Organoids

Relatori

Prof. KRISTEN MARIKO MEIBURGER

Prof. MASSIMO SALVI

Ing. FRANCESCO BRANCIFORTI

Candidato

FILIPPO D'AGOSTINO

2023/2024

Sommario

I modelli in vitro basati su organoidi derivati da cellule staminali rappresentano uno strumento promettente per la modellazione dei tessuti umani. Questi organoidi presentano una somiglianza funzionale e strutturale con gli organi umani, superiore ai tradizionali saggi basati su colture 2D. Sebbene l'analisi dello sviluppo strutturale degli organoidi possa fornire informazioni preziose, l'applicazione di strumenti tradizionali per l'analisi delle immagini risulta problematica a causa dell'eterogeneità e della natura complessa delle morfologie degli organoidi.

D'altra parte, l'*Optical Coherence Tomography* (OCT) rappresenta un metodo di imaging completamente non invasivo e non distruttivo, che non richiede l'impiego di marcatori fluorescenti. Grazie alla sua natura non distruttiva, l'OCT si configura come una soluzione ideale per monitorare la crescita degli organoidi nel tempo e valutare l'effetto di potenziali trattamenti.

L'attenzione principale sarà rivolta all'analisi e all'organizzazione dei dati OCT disponibili. L'obiettivo di questo studio è lo sviluppo di un algoritmo di deep learning per la segmentazione degli organoidi e un algoritmo di tracciamento automatico per monitorarne la crescita e lo sviluppo, estraendo le caratteristiche principali di tutti gli organoidi completamente tracciati. Infine, il framework proposto sarà sottoposto a test e si indagherà la potenziale differenziazione tra gli organoidi che hanno ricevuto un trattamento e quelli che non l'hanno ricevuto.

Indice

Elenco delle figure	IV
Elenco delle tabelle	VI
1 Introduzione	1
1.1 Colture 3D: organoidi	2
1.2 Tecniche di Imaging	4
1.2.1 Bright-Field Microscopy	4
1.2.2 Fluorescence Microscopy	5
1.2.3 Laser Scanning Confocal Microscopy	6
1.2.4 Vantaggi e svantaggi	7
1.3 Optical Coherence Tomography	7
1.4 Segmentazione di immagini mediche	9
1.4.1 Deep Learning	11
1.4.2 Convolutional Neural Network	12
2 Materiali	14
3 Metodi	17
3.1 Preprocessing	17
3.2 CNN K-Net SWIN Transformer	19
3.3 Postprocessing	20
3.4 Metriche di valutazione	22
3.4.1 Dice score	22
3.4.2 Recall	23
3.4.3 Precision	23
3.4.4 Valutazioni	24
3.5 Algoritmo di Tracking	25
3.5.1 Definizione parametri	26
3.5.2 Tracking a doppio ramo	27
3.5.3 Gestione Risultati	33

4 Risultati	36
4.1 Visualizzazione risultati	36
4.2 Analisi di coerenza	38
4.3 Crescita volumetrica	39
5 Conclusione	43
Bibliografia	46

Elenco delle figure

1.1	Workflow per la coltura e lo screening degli organoidi.	3
1.2	Acquisizione tramite Bright-Field Microscopy di organoidi tumorali.	5
1.3	Fluorescence Microscopy di cellule tumorali.	6
1.4	LSCM di organoidi 3D dopo 72 h, 96 h e 120 h di coltivazione.	6
1.5	Esempio di volume OCT con le sue diverse prospettive.	8
1.6	Sistema di scansione OCT.	9
1.7	Esempio di segmentazione di un tumore cerebrale.	10
1.8	Modello di ANN.	11
1.9	Esempio di architettura CNN.	13
2.1	Maximum Intensity Projections (MIPs) lungo la direzione X in differenti istanti temporali.	16
3.1	Pipeline algoritmo	17
3.2	(A) Istogrammi originali. (B) Istogramma di riferimento. (C) Istogrammi normalizzati. (D) Confronti	18
3.3	(A) Immagine originale. (B) Normalizzazione Istogramma. (C) Histogram stretching. (D) Sharpening. (E) Filtro Mediano	19
3.4	Panoramica dell'architettura K-Net utilizzata per la segmentazione degli organoidi.	21
3.5	Risultati intermedi del postprocessing.	22
3.6	Esempio di predizione di un'immagine 2D con il relativo Ground Truth.	23
3.7	A) Andamento del Dice score nei diversi giorni di acquisizione. B) Confronto tra maschere manuali e maschere automatiche.	25
3.8	Visualizzazione dell'algoritmo di tracking.	28
3.9	Andamento del numero degli organoidi nei diversi giorni di acquisizione.	29
3.10	Data Cleaning della matrice con i Dice score per l'analisi delle fusioni.	30
3.11	Fase di assegnazione e analisi fusioni.	31
3.12	Data Cleaning della matrice con i Dice score per l'analisi delle divisioni.	32
3.13	Esempio andamento fusioni/divisioni organoide 113 del sample w8.	32

3.14	Esempio di analisi di coerenza, organoide ID:200, sample: w8	34
4.1	Visualizzazione 3D dei risultati del tracking ottenuti del volume w8.	37
4.2	MIPs relative ai giorni d3, d5, d7 e d9 del campione w8.	38
4.3	Valutazione coerenza organoidi 104, 105, 142.1, 142.2.	39
4.4	Valutazione coerenza e sviluppo dell'organoide 113.	40
4.5	A) Trend volume medio organoidi "Treated". B) Trend volume medio organoidi 'Untreated'.	41
4.6	Comparazione ultimo giorno di Tracking.	42

Elenco delle tabelle

1.1	Confronto tra le caratteristiche delle colture 2D e 3D.	4
1.2	Vantaggi, svantaggi e applicazioni OCT.	9
1.3	Vantaggi e svantaggi CNN.	13
2.1	Divisione dataset in training, validation e test set.	15
2.2	Nuovo Dataset utilizzato per la fase di tracking.	15
3.1	Metriche calcolate sul test set relative al volume p1 w1. (BA) Baseline. (PP) Preprocessing. (PP+RS0+S3D) Preprocessing+postprocessing	24
3.2	Tabella riassuntiva delle metriche di valutazione.	25
3.3	Parametri utilizzati per il tracking	26
3.4	Dataset utilizzato per la fase di tracking.	27
4.1	Numero di organoidi completamenti tracciati per ogni volume.	36
4.2	Volume medio organoidi campioni "Treated"	40
4.3	Volume medio organoidi campioni 'Untreated'	41

Capitolo 1

Introduzione

Nell'arco degli ultimi anni, sono stati fatti progressi notevoli sulla ricerca oncologica andando ad approfondire numerosi aspetti fisiologici del microambiente tumorale. Modelli in vitro e in vivo sono strumenti importanti nella ricerca medica soprattutto per quando riguarda la ricerca sul cancro, consentono di: individuare possibili agenti cancerogeni, sviluppare nuove terapie antitumorali, verificare l'efficacia dei farmaci e di acquisire nuove informazioni riguardo i meccanismi di crescita e metastasi tumorali. Lo studio del corpo umano e i suoi meccanismi a livello cellulare rappresentano ancora un'area di ricerca da approfondire. Le colture di cellule sono un sistema di modello preclinico indispensabile poiché essi consentono la creazione di modelli in vitro utili per una migliore comprensione dei fenomeni biologici. Le cellule all'interno del corpo sono continuamente soggette a segnali e interazioni cellula-cellula per stabilire, mantenere e regolare il loro sviluppo e le loro proprietà biochimiche e fisiologiche. Convenzionalmente vengono utilizzate le colture cellulari in 2D in cui le cellule aderiscono sopra una superficie piatta di vetro/plastica formando strutture monodimensionali. Purtroppo, questo approccio presenta anche degli svantaggi come:

- **Non riflette pienamente l'ambiente cellulare naturale:** all'interno del corpo umano, le cellule interagiscono in un contesto tridimensionale. La crescita su una superficie piatta non rispecchia adeguatamente tali interazioni, rendendo questo metodo inadeguato per lo studio delle funzioni cellulari.
- **Mancanza di predittività:** i test condotti sulle colture 2D spesso non riescono a prevedere con precisione l'efficacia dei farmaci o il risultato delle sperimentazioni cliniche. Questo porta ad un aumento dei costi e dei tassi di insuccesso nel processo di scoperta di nuovi farmaci e nello sviluppo clinico.
- **Problemi derivanti dalla crescita cellulare:** all'interno delle colture 2D standard, i nutrienti sono impiegati per favorire la crescita cellulare,

con conseguente produzione di rifiuti metabolici. Tale processo può causare l'accumulo di sostanze di scarto nocive, la presenza di cellule morte, una riduzione della disponibilità di nutrienti e danni all'ambiente cellulare.

Nonostante questi difetti, le colture 2D continuano ad essere utilizzate perché hanno costi bassi, hanno bisogno poca manutenzione e la presenza di una vasta letteratura comparativa facilita il confronto tra i risultati attuali e quelli degli studi precedenti. Le colture 2D non riescono ad emulare le principali caratteristiche delle cellule perché, a causa dell'appiattimento sulla superficie del supporto, perdono la loro struttura gerarchica e la loro morfologia, questo fenomeno implica la perdita, da parte delle cellule, della loro capacità di riprodursi e differenziarsi. Perciò, per avere una migliore rappresentazione della realtà, c'è la necessità di creare strutture tridimensionali che siano in grado riprodurre fedelmente la complessa struttura e funzionalità degli organi umani, per questo sono stati creati dei modelli più innovativi chiamati organoidi.

1.1 Colture 3D: organoidi

Un organoide è un tessuto 3D auto-organizzato, tipicamente derivato da cellule staminali, che imita la complessità funzionale, strutturale e biologica di un organo. Queste strutture rivestono una crescente importanza nella ricerca oncologica, neurobiologica, sulle cellule staminali e nella scoperta farmacologica, poiché consentono una più accurata modellazione dei tessuti umani. Gli organoidi possono essere ottenuti a partire da *cellule staminali embrionali* (ESCs), *cellule staminali pluripotenti indotte* (iPSCs) e *cellule staminali adulte* [1]. Queste cellule, all'interno di un ambiente extracellulare controllato come matrigel o collagene, possono proliferare e autoassemblarsi formando strutture complesse che riescono a riprodurre la morfologia e le funzionalità di organi e tessuti. Le *cellule staminali* (SC) sono cellule indifferenziate o parzialmente differenziate che sono in grado di auto-rinnovarsi a lungo termine. Queste cellule sono fondamentali per tutte le strutture nel nostro corpo perché sono responsabile del mantenimento dell'omeostasi dei tessuti adulti. Le iPSCs possono differenziarsi in molteplici tipi di cellule mentre, per quanto riguarda le ASCs, si possono differenziare solo nelle cellule presenti nell'organo nella quale si sono originati. Grazie a questa loro caratteristica, organoidi realizzati attraverso le ASCs, riescono ad emulare meglio la capacità dei tessuti di rigenerarsi rispetto alle colture di iPSCs, rendendo queste colture più stabili nel tempo. Tuttavia, le ASCs sono meno utilizzate a causa della loro limitata capacità di differenziarsi e la scarsa reperibilità di campioni per questo, si preferisce l'utilizzo delle iPSCs dato che possono essere ottenute attraverso l'estrazione di fibroblasti, questa caratteristica consente anche di poter creare colture specifiche per diversi pazienti nella medicina di precisione [1]. Attraverso l'utilizzo di organoidi derivati

da iPSCs, rispetto alle ASCs, è possibile ottenere tessuti più simili allo stato fetale. Questo permette di sviluppare modelli di studio per comprendere meglio i processi di organogenesi e le patologie ad essa correlate, con questo approccio sono stati sviluppati diversi organoidi che emulano con successo diversi tessuti come: cervello, occhi, reni e polmoni [1]. Partendo da cellule iPSCS isolate, esse vengono inizialmente coltivate in colture 2D con l'obiettivo di promuoverne la crescita. In seguito, le cellule vengono incorporate nel Matrigel e successivamente questa miscela viene suddivisa in piccole goccioline, che sono poi inserite all'interno di un'incubatrice. Qui, avviene il processo di solidificazione, formando una sorta di "cupola" in cui le cellule, con il supporto adeguato di fattori di crescita, continuano il loro sviluppo e la loro differenziazione, dando origine a tessuti specifici tridimensionali. Nelle fasi successive, attraverso l'impiego di tecniche di imaging, è possibile monitorare e visualizzare la crescita e la differenziazione degli organoidi. Si può procedere con la ricostruzione in 3D delle strutture, l'analisi dettagliata della loro composizione, morfologia e vitalità cellulare, nonché la valutazione della differenziazione cellulare stessa [Fig 1.1] [2]. Pertanto, la coltura 3D può costituire uno strumento fonamen-

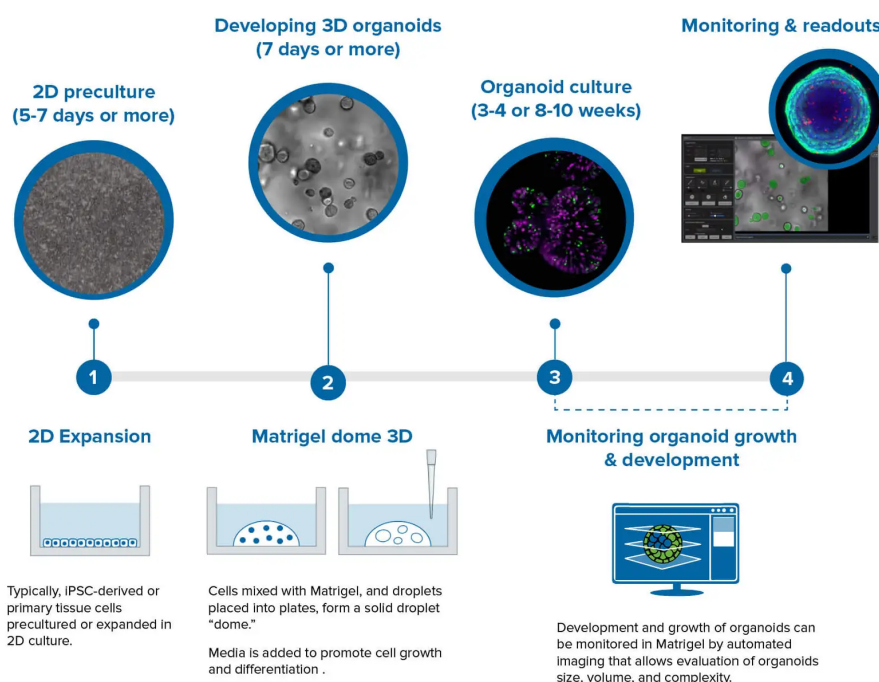


Figura 1.1: Workflow per la coltura e lo screening degli organoidi.

tale in quanto riflettono più accuratamente i microambienti tissutali, le interazioni cellula-cellula e i processi biologici che si verificano in vivo. La Tabella 1.1 riassume le differenze principali dei modelli di coltura 2D e 3D [3].

Confronto tra i modelli 2D e 3D		
Caratteristiche	2D	3D
Costo	Basso	Costoso
Differenziazione	Scarsa	Ben differenziati
Espressione genica	Diversa dai modelli in vivo	Simile ai modelli in vivo
Forma cellule	Piatte	Struttura naturale 3D
Interazioni cellula-cellula	Limitata	Simile ai modelli in vivo
Riproducibilità esperimento	Alta	Bassa
Supporto	Piastra di Petri	Matrice Extracellulare
Tempi	Minuti/Ore	Ore/Giorni

Tabella 1.1: Confronto tra le caratteristiche delle colture 2D e 3D.

Nel paragrafo successivo, esploreremo le diverse tecniche di imaging impiegate per l'acquisizione di immagini delle colture di organoidi utilizzate per la ricostruzione dei loro volumi 3D.

1.2 Tecniche di Imaging

Gli organoidi, possono essere considerati modelli semplificati degli organi in vivo, offrono una rappresentazione più accurata degli organi rispetto alle colture 2D. Essi emulano più fedelmente l'espressione genica, l'architettura tissutale e le funzioni metaboliche dell'organo che rappresentano. Con l'evoluzione delle tecniche di coltura 3D, emerge la necessità di perfezionare le tecniche di imaging. Le immagini 2D si rivelano insufficienti per un'analisi accurata, in quanto non sono in grado di rappresentare adeguatamente la complessa struttura tridimensionale degli organoidi. Mediante l'utilizzo delle immagini acquisite, è possibile procedere alla ricostruzione del volume in tre dimensioni. Questi dati consentono di effettuare due tipi di analisi:

- **Analisi quantitativa:** che comprende il conteggio delle cellule e dei loro componenti chimici.
- **Analisi morfologica:** che include misurazioni come l'area, il perimetro, il diametro e la sfericità di tutti gli organoidi.

Di seguito vengono presentate alcune delle tecniche di imaging comunemente utilizzate per l'acquisizione di immagini relative alle colture di organoidi.

1.2.1 Bright-Field Microscopy

La *Bright-Field Microscopy* è una tecnica impiegata per l'analisi morfologica degli organoidi e per la determinazione di parametri bidimensionali quali la lunghezza e

l'area. Questa metodologia non richiede alcuna preparazione specifica dei campioni. Parametri quali il diametro, il perimetro e l'area possono essere utilizzati come indicatori per valutare la crescita degli organoidi e determinare il loro stato vitale o non vitale. Questi dati possono fornire informazioni preziose sulla dinamica di crescita e sulla salute degli organoidi [Figura 1.2] [4].

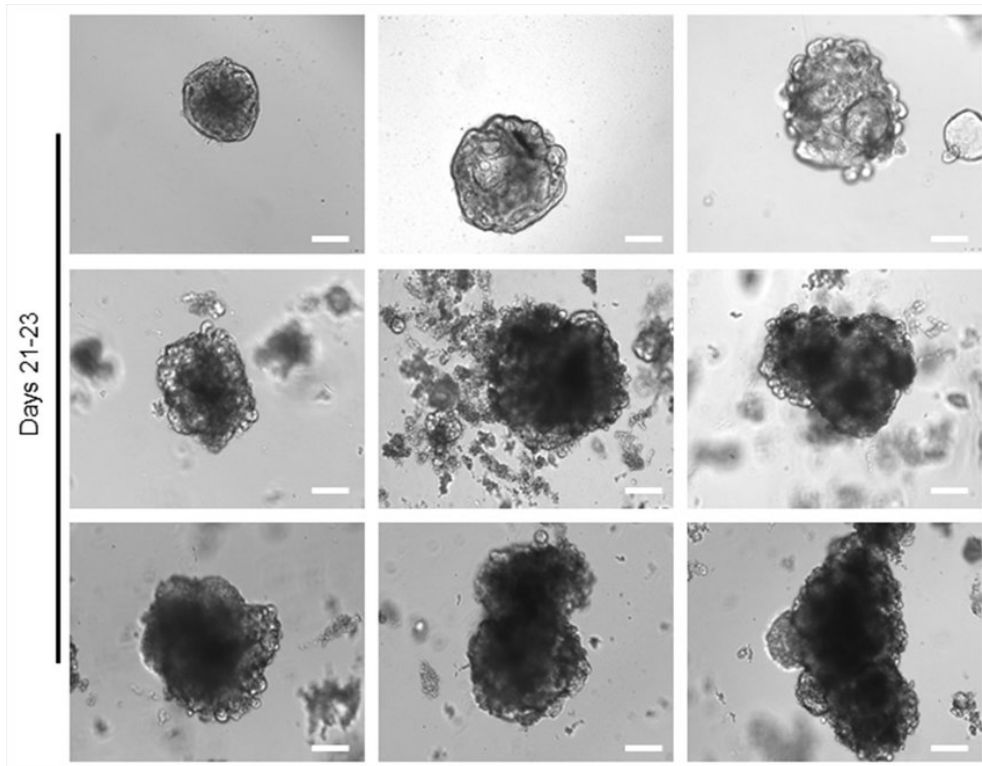


Figura 1.2: Acquisizione tramite Bright-Field Microscopy di organoidi tumorali.

1.2.2 Fluorescence Microscopy

La *Fluorescence Microscopy*, mediante l'impiego di una specifica lunghezza d'onda come sorgente luminosa, permette di irradiare il campione inducendolo a emettere fluorescenza. Questo consente l'analisi della forma e della posizione dell'organoide. Sostanze intracellulari, come l'ATP o il NADPH, possono emettere fluorescenza in seguito a irradiazione ultravioletta, rendendo possibile la visualizzazione e la quantificazione di specifiche molecole. Il principale vantaggio dell'imaging a fluorescenza risiede nella sua capacità di visualizzare con grande specificità tessuti, tipi cellulari e particolari biomolecole, grazie all'introduzione di sonde e marcatori fluorescenti. Questa tecnica offre quindi un potente strumento per lo studio dettagliato degli organoidi [Figura 1.3] [4].

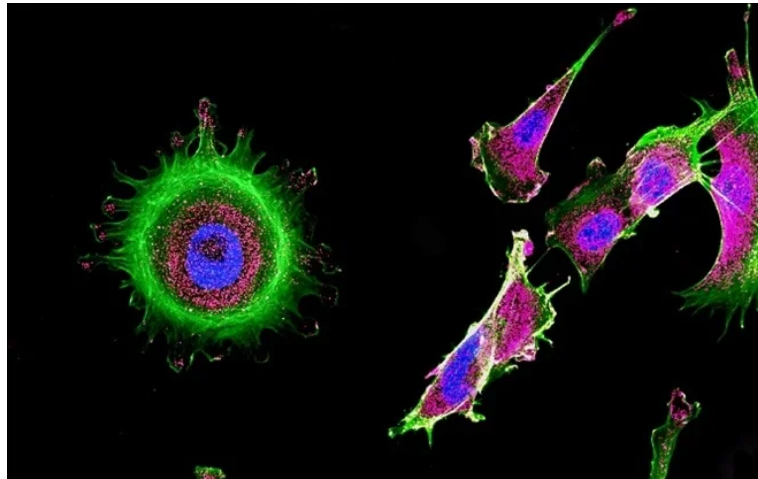


Figura 1.3: Fluorescence Microscopy di cellule tumorali.

1.2.3 Laser Scanning Confocal Microscopy

Il *Laser Scanning Confocal Microscopy* (LSCM) rappresenta una delle metodologie più impiegate nell'ambito dell'imaging di organoidi. Questa tecnica, mediante l'utilizzo di un laser, esegue una scansione dettagliata del campione, analizzandolo punto per punto. Da questo processo, si ottiene un volume tridimensionale dato dall'unione di tutte le sezioni acquisite. L'analisi di questi dati consente la quantificazione di una serie di parametri, tra cui la morfologia, l'area, l'attività biochimica e il conteggio delle cellule. Queste informazioni forniscono un quadro dettagliato della crescita e dello sviluppo degli organoidi [Figura 1.4] [4].

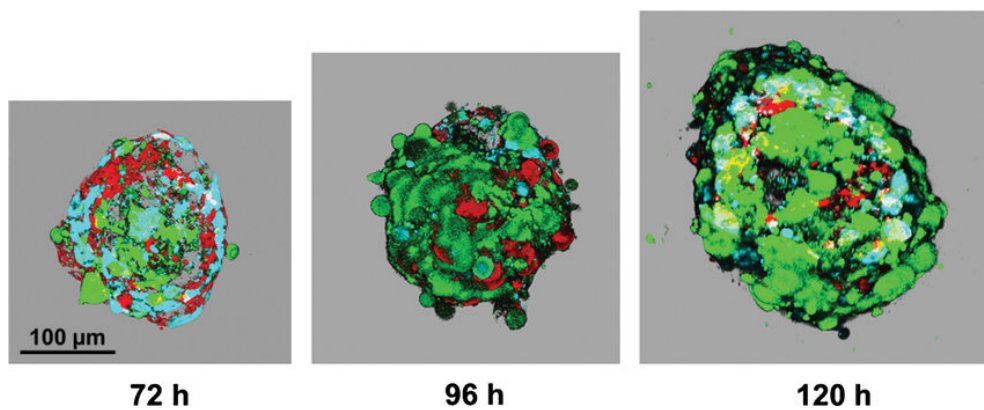


Figura 1.4: LSCM di organoidi 3D dopo 72 h, 96 h e 120 h di coltivazione.

1.2.4 Vantaggi e svantaggi

Gli organoidi, strutture tridimensionali che simulano le funzioni degli organi biologici, hanno trovato numerose applicazioni nel campo della ricerca biomedica negli ultimi anni. La loro struttura complessa richiede l'impiego di tecnologie di imaging avanzate, capaci di valutare l'organoide nel suo insieme, minimizzando le interferenze dovute a fonti esterne. È essenziale visualizzare la struttura degli organoidi per monitorare la loro crescita e sviluppo. Questo approccio consente una valutazione accurata del loro potenziale e delle loro limitazioni nell'ambito della ricerca biomedica. Tuttavia, ogni tecnologia di imaging presenta dei limiti. Ad esempio, la *Bright-Field Microscopy*, sebbene sia molto utile per le misure di lunghezza e aree, permette solo misure parziali della forma, poiché le informazioni tridimensionali non vengono catturate [4]. Un altro problema è che la qualità dell'osservazione si riduce quando il contrasto è troppo elevato, creando distorsioni nell'immagine. La *Fluorescence Microscopy* ha una bassa risoluzione e non è in grado di effettuare sezioni ottiche (acquisire immagini bidimensionali a diversi livelli di profondità), rendendola inadatta per l'imaging tridimensionale degli organoidi [4]. D'altra parte, la LSCM può essere utilizzata per l'imaging tridimensionale degli organoidi, ma la sua bassa velocità di scansione limita le acquisizioni a lungo termine a causa delle interferenze esterne [4]. Tutte le metodologie di imaging finora descritte presentano limitazioni per l'acquisizione appropriata dei dati necessari. Nella sezione successiva, verrà introdotta la metodologia di imaging impiegata in questo lavoro di tesi: l'*Optical Coherence Tomography* (OCT). Questa tecnica avanzata offre una serie di vantaggi significativi rispetto alle metodologie tradizionali, permettendo l'acquisizione di dati più accurati e dettagliati.

1.3 Optical Coherence Tomography

L'*Optical Coherence Tomography* (OCT) è una tecnologia che ha trovato ampio impiego nella ricerca biomedica e clinica. Questa tecnica, basata sull'interferenza della luce retrodiffusa, è in grado di ricostruire un campione in tre dimensioni (3D) con alta risoluzione e velocità. L'OCT può fornire una risoluzione su scala micrometrica e una profondità di scansione di circa 1-3 mm per l'imaging dei tessuti. Questa tecnologia migliora la capacità di visualizzare in 3D i singoli organoidi, fornisce una visione migliore della loro struttura interna e permette di rivelare caratteristiche anatomiche non osservabili con altre tecnologie [5]. Inoltre, l'OCT è stata utilizzata per monitorare lo sviluppo delle cellule, consentendone la visualizzazione e la quantificazione delle dinamiche tissutali. Con il continuo miglioramento e l'ottimizzazione di questa tecnologia, l'OCT viene frequentemente utilizzato nello studio di farmaci o per valutare l'efficacia di specifici trattamenti. Questa tecnologia si avvale dell'*interferometria*, un metodo che sfrutta l'interferenza

tra onde coerenti per misurare la riflessione di onde luminose da diverse strutture biologiche [6]. La luce, elemento chiave di questa tecnica, è caratterizzata da una specifica ampiezza e lunghezza d'onda, dovute alla sua natura ondulatoria. Quando il fascio luminoso interagisce con l'organoide, induce un'oscillazione alla stessa frequenza dei fotoni incidenti. Questa oscillazione provoca a sua volta l'emissione di fotoni, che vengono rilevati da un interferometro. L'interferometro elabora il segnale e genera l'immagine in cui vengono rappresentate in maniera dettagliata le strutture interne [Figura 1.5] [6].

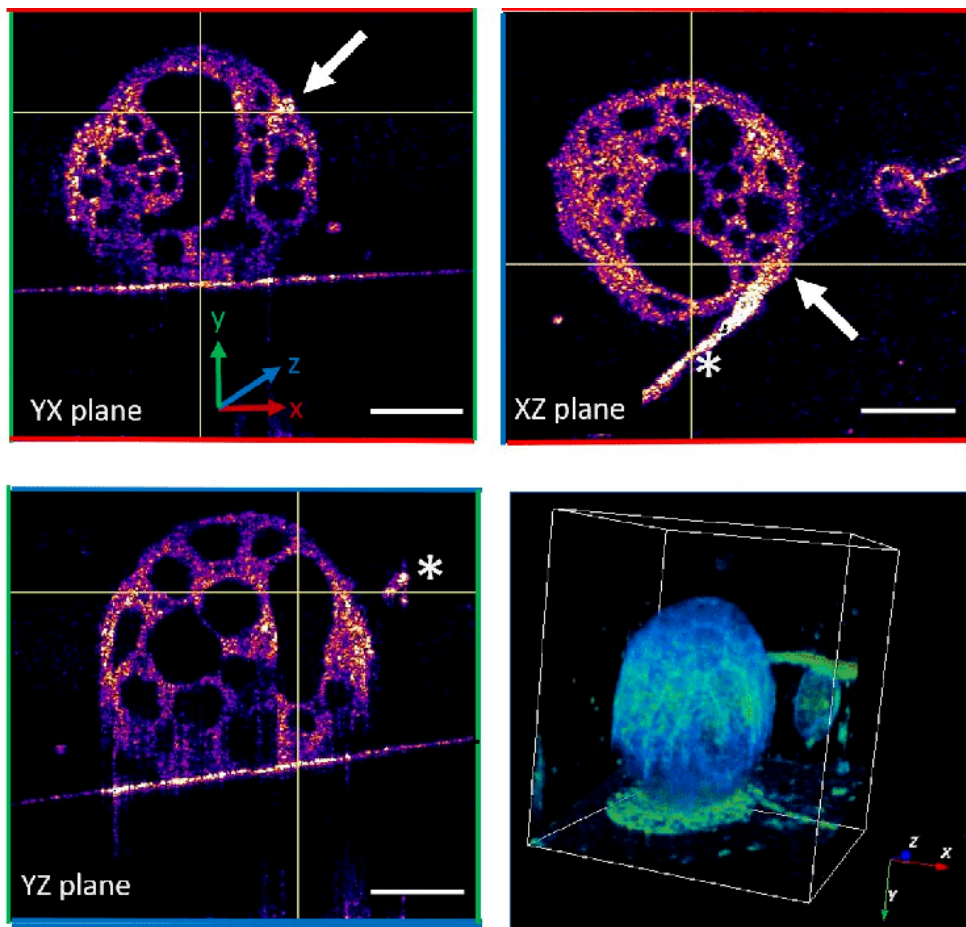


Figura 1.5: Esempio di volume OCT con le sue diverse prospettive.

Nel dettaglio [Figura 1.6], la luce emessa dalla sorgente viene suddivisa da un divisore di fascio in due percorsi distinti, denominati braccio di riferimento e braccio campione. La combinazione delle luci riflesse dai due bracci sul rivelatore genera fenomeni di interferenza, manifestati come variazione rapida dell'intensità luminosa, i quali sono rilevabili solo se i tempi di percorrenza dei due bracci sono comparabili

[6]. La presenza di interferenze fornisce un indicatore relativo della distanza percorsa dalla luce. Attraverso l'impiego di un braccio di scansione per mappare il campione nelle coordinate x e y e sfruttando l'interferometria per acquisire informazioni sulla profondità, è possibile ottenere un'immagine tridimensionale completa del campione.

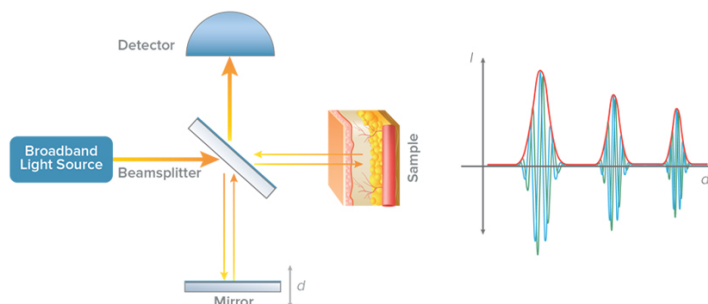


Figura 1.6: Sistema di scansione OCT.

La Tabella 1.2 riassume i principali svantaggi e vantaggi di questa tecnica e le sue applicazioni.

Optical Coherence Tomography		
Vantaggi	Svantaggi	Applicazioni
Non invasivo	Acquisizioni limitate dovute alla presenza di vasi sanguigni.	Volumi 3D
I campioni non devono essere trattati	Limitato potere di penetrazione	Visualizzazione e quantificazione dei parametri cellulari
Tracking a lungo termine	Non adatta per campioni trasparenti	Monitoraggio crescita organoidi
Alta velocità di scansione risoluzione		

Tabella 1.2: Vantaggi, svantaggi e applicazioni OCT.

Attraverso l'impiego dell'*Optical Coherence Tomography*, è possibile acquisire immagini di alto contrasto attraverso le quali è possibile ricostruire i volumi tridimensionali che mettono in evidenza le strutture interne degli organoidi. I volumi ottenuti trovano generalmente applicazione nel campo della segmentazione medica. Questo processo, in combinazione con avanzati modelli di *Intelligenza Artificiale*, costituisce un elemento chiave per le attività diagnostiche fornendo dati di grande importanza per la ricerca scientifica.

1.4 Segmentazione di immagini mediche

Le tecniche di imaging medico, che permettono la generazione di immagini dettagliate degli organi interni del corpo umano, sono diventate strumenti fondamentali

per la diagnosi, l'esame e la pianificazione dei trattamenti in ambito medico. Da quando queste tecnologie sono state introdotte, la loro importanza nella pratica medica è cresciuta in modo esponenziale. L'analisi accurata ed efficiente di queste immagini è essenziale per sfruttare al meglio le informazioni cliniche che esse contengono. Tuttavia, la semplice osservazione può non essere sufficiente per una diagnosi accurata o per la pianificazione di interventi chirurgici o di radioterapia. Pertanto, la segmentazione delle immagini, un processo che implica l'utilizzo di algoritmi per delineare le strutture anatomiche e altre regioni di interesse (ROI) [Figura 1.7], è uno dei compiti più impegnativi nell'analisi delle immagini mediche. Questo processo identifica i pixel che appartengono agli organi o alle lesioni rispetto allo sfondo, fornendo informazioni critiche sulle forme e sui volumi di questi organi. La segmentazione delle immagini è cruciale per ottenere informazioni critiche sulle forme e sui volumi degli organi, il che è essenziale per una diagnosi accurata. Questo compito, sebbene impegnativo, è fondamentale per sfruttare appieno le informazioni cliniche contenute nelle immagini mediche.

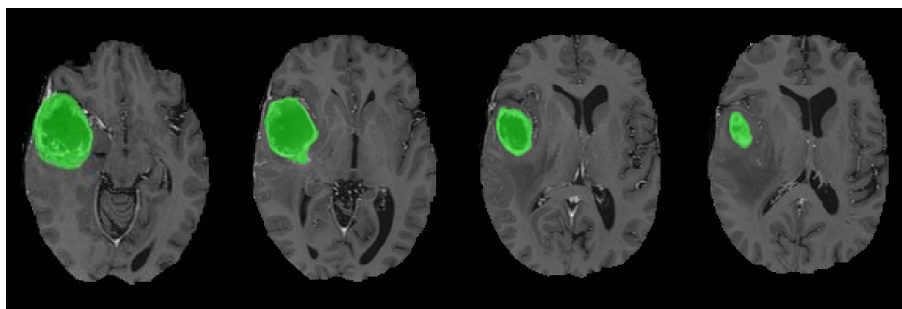


Figura 1.7: Esempio di segmentazione di un tumore cerebrale.

Negli ultimi anni, l'evoluzione delle tecniche di segmentazione delle immagini ha raggiunto nuovi livelli di complessità e precisione, grazie all'introduzione di architetture avanzate di reti neurali convoluzionali. Nonostante l'avvento di questi metodi innovativi, esistono numerose tecniche di segmentazione che possono essere raggruppate in due categorie principali: tecniche di segmentazione tradizionali e tecniche basate sul deep learning. Le tecniche di segmentazione tradizionali, che includono metodi come la sogliatura, la crescita di regioni e le tecniche basate sui bordi, sono state a lungo la colonna portante dell'analisi delle immagini. D'altra parte, le tecniche di deep learning, che sfruttano la potenza delle reti neurali convoluzionali, stanno rivoluzionando il campo, offrendo un livello di precisione e di dettaglio senza precedenti. Nel paragrafo seguente, presenteremo una panoramica generale delle tecniche basate sul deep learning, esplorando i loro punti di forza, le loro limitazioni e le potenziali applicazioni in vari ambiti della ricerca e della pratica clinica.

1.4.1 Deep Learning

La segmentazione delle immagini mediche riveste un ruolo fondamentale nell'ambito della diagnostica e del trattamento medico, poiché mira a trasformare un'immagine di input in una rappresentazione significativa per l'analisi anatomica. Questo processo è cruciale per identificare la regione di interesse (ROI), determinare le sue dimensioni e assistere il medico nella decisione riguardante la terapia da applicare al paziente. L'efficacia della segmentazione delle immagini si riflette nell'agevolare l'analisi delle immagini mediche, mettendo in evidenza in modo accurato la ROI, un passo fondamentale per una diagnosi e un trattamento più accurato. La segmentazione è resa difficile a causa delle varie restrizioni delle tecniche di acquisizione delle immagini mediche, dal tipo di patologia e dalle diverse variazioni biologiche. Negli ultimi anni, le reti di apprendimento profondo hanno contribuito allo sviluppo di nuovi modelli di segmentazione delle immagini con un miglioramento delle prestazioni raggiungendo alti tassi di accuratezza. Le tecniche di segmentazione delle immagini possono essere suddivise in due categorie principali: segmentazione semantica e segmentazione di istanza. La segmentazione semantica può essere vista come un problema di classificazione dei pixel, in cui ogni pixel dell'immagine viene assegnato a una classe specifica. D'altra parte, la segmentazione di istanza localizza e delinea ogni oggetto di interesse presente nell'immagine di input. Il *Deep Learning* è uno degli approcci fondamentali dell'intelligenza artificiale. Gli algoritmi di *Deep Learning* utilizzano vari layers per costruire una *artificial neural network* (ANN). Una ANN è costituita da un layer di ingresso, vari layers nascosti e un layer di uscita. Il layer di ingresso riceve il segnale di input, mentre il layer di uscita prende decisioni basate su tale input. I layers nascosti, situati tra il layer di ingresso e quello di uscita, si occupano di eseguire i calcoli, elaborare e apprendere modelli complessi grazie ai dati di input [7].

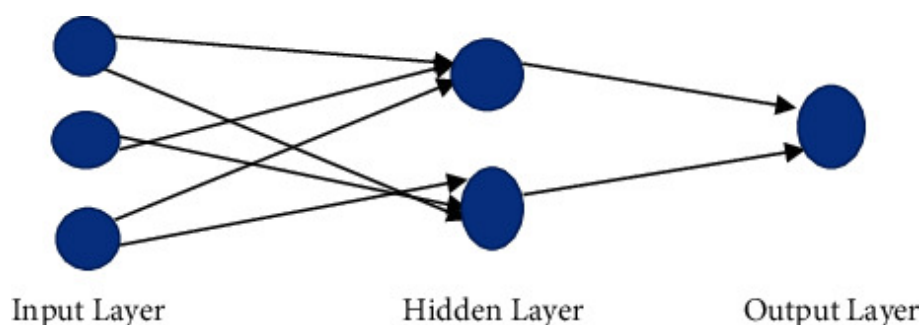


Figura 1.8: Modello di ANN.

1.4.2 Convolutional Neural Network

Una *Convolutional Neural Network* (CNN) è un algoritmo di *Deep Learning* che prende in input un'immagine e assegna un'importanza (pesi) ai vari aspetti o oggetti presenti nell'immagine, permettendo di distinguerli tra loro. Questo tipo di rete può essere composto da decine o centinaia di strati, ognuno dei quali è in grado di rilevare diverse caratteristiche di un'immagine. Durante il processo di addestramento, a ciascuna immagine vengono applicati filtri a diverse risoluzioni e l'output di ogni immagine convoluta viene utilizzato come input per lo strato successivo. I filtri possono rilevare inizialmente caratteristiche molto semplici, come la luminosità o i bordi, e progressivamente diventare più complessi fino a includere caratteristiche che definiscono in modo univoco l'oggetto. La potenza delle CNN risiede nella loro capacità di apprendere automaticamente e adattarsi a queste caratteristiche, permettendo una classificazione accurata e robusta delle immagini. Questo le rende uno strumento fondamentale in molte applicazioni di *Computer Vision* e *Image Processing*.

Una CNN è composta da tre layers principali:

- **Convolutional layer:** è in grado di identificare caratteristiche distintive, come bordi o altri elementi visivi, all'interno di un'immagine. Il livello di convoluzione della rete esegue un'operazione matematica che consiste nella moltiplicazione dei pixel vicini di un'immagine con i kernel. La CNN impiega una varietà di kernel per convolvere l'immagine fornita e generare le sue mappe di caratteristiche. Questo processo consente alla rete di apprendere e riconoscere schemi complessi all'interno delle immagini [8].
- **Pooling layer:** ha il compito di ridurre le dimensioni spaziali (larghezza e altezza) dei dati di ingresso per gli strati successivi della rete. Questa operazione, nota come sottocampionamento, non altera la profondità dei dati. La riduzione delle dimensioni ottenuta attraverso il sottocampionamento contribuisce a diminuire i requisiti computazionali per i livelli successivi della rete, rendendo il modello più efficiente dal punto di vista computazionale [8].
- **Fully-connected layer:** i valori dei pixel dell'immagine di input non sono direttamente connessi ai nodi del livello di uscita nei livelli parzialmente connessi. Tuttavia, nel Fully-connected layer, ogni nodo del livello di uscita è direttamente associato a un nodo del livello precedente. Questo strato svolge un ruolo cruciale nella classificazione, poiché si basa sulle caratteristiche estratte dagli strati precedenti e dai loro vari filtri. Questo processo permette una maggiore precisione nell'identificazione e nella classificazione dei dati dell'immagine [8].

In Figura 1.9 viene rappresentato un esempio dell'architettura di una rete di Deep Learning CNN.

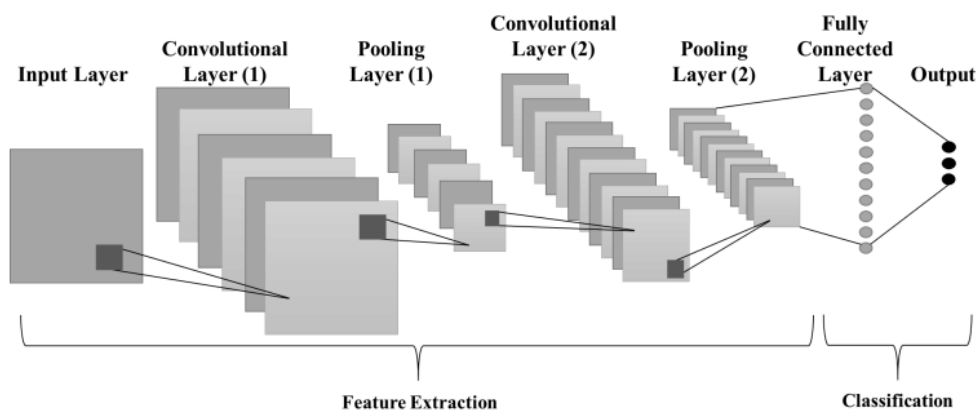


Figura 1.9: Esempio di architettura CNN.

La Tabella 1.3 riassume i principali svantaggi e vantaggi dell'utilizzo di questi modelli di intelligenza artificiale [8].

<i>Convolutional Neural Network</i>	
Vantaggi	Svantaggi
Elaborazione efficiente delle immagini Elevata accuratezza Robusto al rumore Allenamento differenziato Estrazione automatica delle caratteristiche	Elevati requisiti computazionali Richiedono grandi dataset Capacità limitata di generalizzazione Mancanza di interpretabilità

Tabella 1.3: Vantaggi e svantaggi CNN.

Nei capitoli successivi, si discuteranno in dettaglio i materiali impiegati, in particolare il dataset. Saranno delineati i vari passaggi adottati per effettuare la segmentazione delle immagini OCT degli organoidi. Inoltre, verrà fornita una spiegazione approfondita del funzionamento dell'algoritmo di tracking sviluppato per monitorare la crescita di tali organoidi durante i diversi giorni di acquisizione.

Capitolo 2

Materiali

L'*Optical Coherence Tomography* (OCT) si sta rivelando una tecnologia di imaging molto promettente per l'analisi degli organoidi. Questa tecnologia offre la possibilità di ottenere immagini tridimensionali ad alta risoluzione dei tessuti biologici. Questa caratteristica è fondamentale per il monitoraggio longitudinale degli organoidi, consentendo un'analisi quantitativa di numerose caratteristiche importanti della crescita e dello stato del singolo organoide, come il volume e la struttura interna. Nonostante il suo potenziale, l'identificazione e la quantificazione accurata degli organoidi nelle immagini OCT rappresentano ancora una sfida. Questo è dovuto a vari fattori, tra cui lo scarso contrasto delle immagini, i confini poco chiari degli organoidi, il rumore di fondo e le piccole dimensioni dei bersagli degli organoidi.

Nel contesto di questo studio, sono stati impiegati due dataset distinti: uno destinato all'addestramento della rete neurale e l'altro per valutare l'efficacia dell'algoritmo di tracking. I campioni contrassegnati con "p1" rappresentano volumi che sono stati sottoposti a un trattamento specifico e sono stati utilizzati per studiare il fenomeno noto come *Drug Tolerant Persister* (DPT). Questo fenomeno riguarda la ricrescita delle cellule tumorali dopo l'esposizione a farmaci antitumorali. L'osservazione di tale fenomeno permette di indagare i meccanismi di riparazione delle cellule e la loro capacità di alterare il DNA per aumentare la loro resistenza ai farmaci.

Il dataset utilizzato per l'addestramento della rete CNN è composto da tre colture di organoidi differenti, ciascuno dei quali contiene diversi volumi tridimensionali acquisiti in momenti temporali differenti. Questi volumi sono stati suddivisi in training, validation e test set. I campioni w2 e w6, che includono sei volumi corrispondenti a diversi giorni di acquisizione, sono stati impiegati nelle fasi di addestramento e validazione della rete. Il campione p1_w1, che presenta un intervallo di acquisizione più esteso rispetto agli altri con un totale di 10 volumi, è stato utilizzato per costituire il test set. Da quest'ultimo sono state successivamente calcolate le metriche di valutazione della rete CNN addestrata. Ogni volume 3D è

composto da 698 slices, ognuna delle quali è formata da 1200x800 pixel.

Subset	Sample	Days	Dimension
Training	w2	d5, d7, d11, d13	1200x800x698
	w6	d5, d7, d9, d11	1200x800x698
Validation	w2	d9	1200x800x698
	w6	d13	1200x800x698
Testing	p1_w1	d1, d3, d5, d7, d9, d11, d13, d15, d17, d21	1200x800x698

Tabella 2.1: Divisione dataset in training, validation e test set.

Il secondo insieme di dati, impiegato per valutare le prestazioni dell’algoritmo di tracking, comprende sette volumi. Ciascuno di questi volumi è contrassegnato come ’Treated’ o ’Untreated’, a seconda se ha subito il trattamento antitumorale o meno. La Tabella 2.2 fornisce una panoramica dei dati relativi a questo insieme, includendo informazioni quali il nome del campione, l’etichetta e le dimensioni dei volumi 3D che contengono gli organoidi.

Sample	Label	Dimension
p1_w6	Treated	1200x800x698
p1_w7	Treated	1200x880x698
p1_w8	Treated	1200x830x698
w4	Untreated	1200x857x698
w5	Untreated	1200x811x698
w7	Untreated	1200x840x698
w8	Untreated	1200x850x698

Tabella 2.2: Nuovo Dataset utilizzato per la fase di tracking.

La Figura 2.1 presenta una sequenza di immagini relative a tre volumi distinti, ognuno proveniente da diverse colture di organoidi. Questi volumi, utilizzati durante le fasi di addestramento della rete, sono stati selezionati in base agli ultimi giorni di acquisizione di ogni volume. Attraverso queste immagini, è possibile avere una rappresentazione visiva della crescita dinamica degli organoidi, permettendo di osservare e documentare i cambiamenti progressivi in termini di dimensioni e struttura durante le fasi di sviluppo degli organoidi.

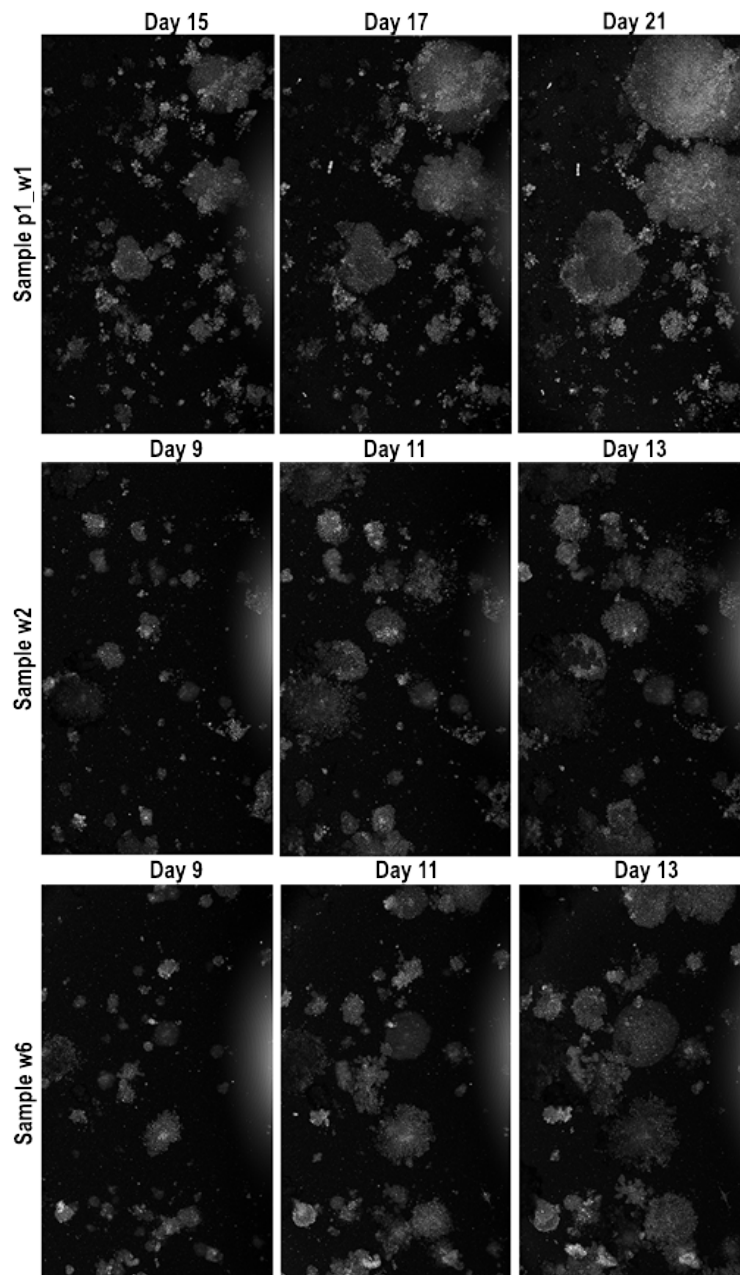


Figura 2.1: Maximum Intensity Projections (MIPs) lungo la direzione X in differenti istanti temporali.

Capitolo 3

Metodi

Nel presente capitolo, si delineano le varie fasi implementate, a partire dall'addestramento della rete neurale fino al monitoraggio della crescita degli organoidi. Il processo ha avuto inizio con il preprocessing del set di immagini, seguito dalla suddivisione del dataset in train, validation e test set. In seguito, verrà descritta l'architettura della rete neurale impiegata per eseguire la segmentazione automatica degli organoidi nelle immagini OCT. Successivamente, con la fase di post-elaborazione, si procede con il calcolo di specifiche metriche per valutare l'accuratezza delle diverse segmentazioni effettuate su tutti i dataset. Dopo aver ottenuto le segmentazioni automatiche dalla rete, si procederà alla fase di tracking utilizzando i volumi a disposizione. La Figura 3.1 rappresenta la pipeline di tutti gli steps effettuati.

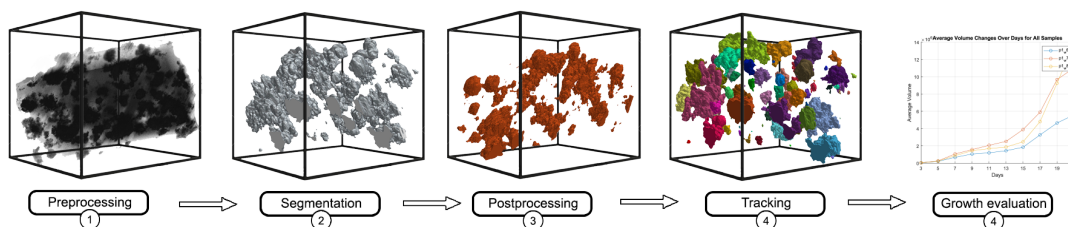


Figura 3.1: Pipeline algoritmo

3.1 Preprocessing

Nel contesto del preprocessing dei dati, il primo passo è stato quello di uniformare la dinamica di intensità del dataset. Questo è stato ottenuto attraverso una normalizzazione rispetto al 95° percentile, calcolato sulla base della somma degli istogrammi dei volumi presenti nel dataset di training [9]. Successivamente, è stato

applicato uno stretching dell'istogramma, compreso tra il 1° e il 99° percentile, su ciascun volume normalizzato. Questo passaggio ha permesso di aumentare il contrasto tra gli organoidi e lo sfondo, rappresentato dalla goccia. La Figura 3.2 offre una visualizzazione dettagliata dell'andamento dei valori dei pixel attraverso i diversi volumi impiegati per l'addestramento della rete neurale. Inoltre, illustra il risultato conseguito mediante la normalizzazione degli istogrammi.

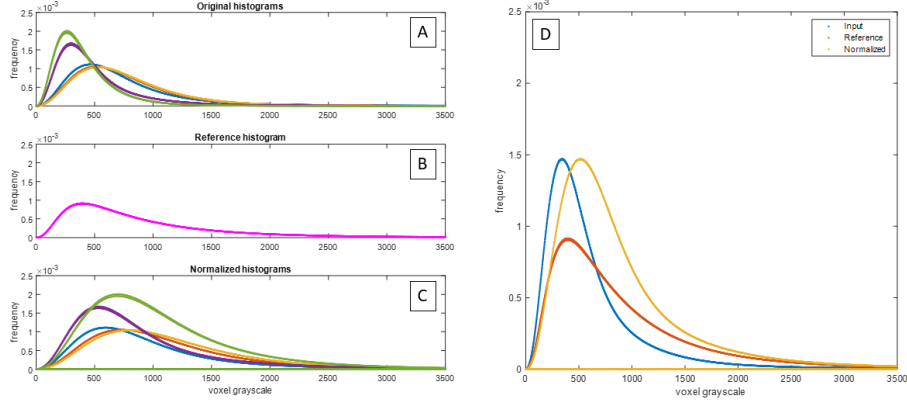


Figura 3.2: (A) Istogrammi originali. (B) Istogramma di riferimento. (C) Istogrammi normalizzati. (D) Confronti

È noto che il rumore speckle, tipicamente presente nelle scansioni OCT, segue una distribuzione di Poisson [10]. Pertanto, è stata operata una ridistribuzione del rumore da una distribuzione Poissoniana a una Gaussiana, attraverso l'applicazione della radice quadrata. Questo passaggio ha migliorato l'efficacia dei successivi step di filtraggio. Infine, è stato eseguito uno sharpening dei volumi sottraendo a ciascuno di essi una versione filtrata (smoothed) dello stesso volume (U), definita “un-sharpening mask”. Questo processo ha permesso di evidenziare i contorni (edges) degli organoidi. La formula utilizzata è stata:

$$V_u(x, y, z) = \frac{b}{2b - 1} V(x, y, z) + \frac{1 - b}{2b - 1} U(x, y, z)$$

Dove il parametro b indica il peso relativo tra il volume originale e la sua controparte filtrata. Nel contesto di questo studio, è stato scelto un valore di b pari a $\frac{2}{3}$. U rappresenta la versione filtrata del volume V , ottenuta mediante l'applicazione del filtro gaussiano $g(x,y,z)$ con parametro σ fissato a 3.

$$g(x, y, z) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2 + z^2}{2\sigma^2}\right)$$

In conclusione, è stato applicato un filtro mediano per la riduzione del rumore, utilizzando un kernel cubico di dimensione 7. Successivamente, i volumi sono stati convertiti in formato uint8.

La Figura 3.3 rappresenta i risultati intermedi ottenuti durante la fase di preprocessing. Questo processo ha rivestito un ruolo cruciale anche per la segmentazione manuale degli organoidi, un passaggio indispensabile per l'addestramento della rete neurale. La segmentazione manuale rappresenta una fase cardine, in quanto l'efficacia della rete neurale dipende direttamente da essa.

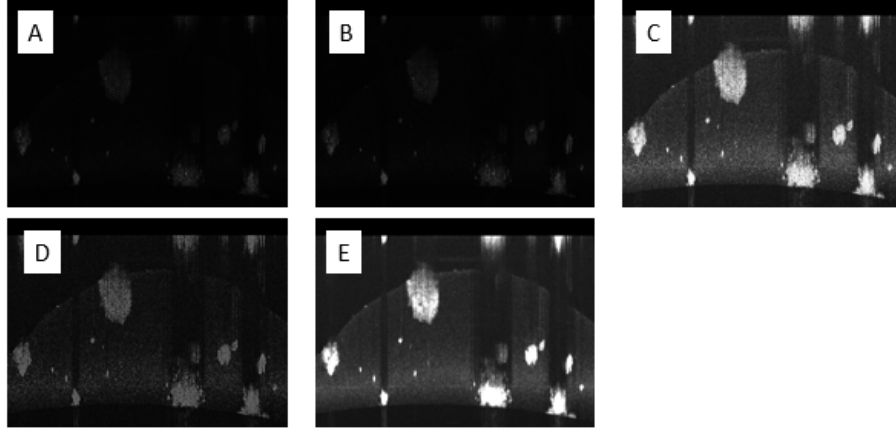


Figura 3.3: (A) Immagine originale. (B) Normalizzazione Istogramma. (C) Histogram stretching. (D) Sharpening. (E) Filtro Mediano

3.2 CNN K-Net SWIN Transformer

Il dataset utilizzato per questo studio consiste in volumi tridimensionali delle colture di organoidi acquisiti tramite OCT. Durante la fase di addestramento della rete CNN, sono state estratte singole sezioni 2D da ciascun volume. Questo approccio ha permesso l'implementazione di un'architettura 2D, ottimizzando la predizione su base slice-per-slice.

Nel contesto della segmentazione degli organoidi, è stata adottata un'architettura di rete chiamata K-Net, una specifica rete CNN utilizzata per questo scopo. L'implementazione di K-Net [10] è stata realizzata attraverso il framework MM Segmentation, un toolbox che facilita l'implementazione e la valutazione della segmentazione semantica [11]. La segmentazione semantica è un processo che raggruppa diverse parti di un'immagine all'interno di specifiche classi di appartenenza. Questo processo è essenzialmente una forma di predizione a livello di pixel, in quanto ogni pixel di un'immagine viene classificato in una categoria specifica. Questa metodologia consente una segmentazione accurata e dettagliata degli organoidi.

Durante la fase iniziale dell'addestramento, i kernel convoluzionali della rete neurale vengono inizializzati in modo casuale. Questi kernel vengono poi aggiornati

iterativamente, in base alle caratteristiche dell'oggetto di segmentazione identificato nell'immagine bidimensionale. Nella fase conclusiva dell'addestramento, i kernel, ora ottimizzati, vengono utilizzati per eseguire operazioni di convoluzione sulle immagini, permettendo così di generare una predizione finale accurata.

Quando si parla di reti neurali, con il termine backbone si intende la capacità da parte della rete di estrarre features dalle immagini in input. In questo studio è stato utilizzato uno SWIN transformer [12], un tipo di backbone che ha permesso di estrarre molteplici caratteristiche dalle immagini in ingresso alla rete. Questa scelta è dovuta alla sua efficienza e alla sua capacità di catturare le dipendenze a lungo raggio nelle immagini 2D. La struttura di decoder è composta da un modulo in grado di generare kernel diversi, con l'obiettivo di catturare le diverse dipendenze globali nelle forme degli organoidi. I kernel convoluzionali sono resi dinamici in quanto dipendono dalle mappe di caratteristiche in ingresso, consentendo l'estrazione di diverse caratteristiche semantiche. Sono stati utilizzati tre blocchi sequenziali per aggiornare i kernel, con ogni blocco che aggiorna i kernel in base alle caratteristiche precedenti. Per migliorare la fase di predizione durante l'addestramento, è stato utilizzato un kernel ausiliario che opera alla risoluzione di $\frac{1}{4}$ rispetto all'input originale, utilizzando lo stesso decoder.

La pre-elaborazione dei dati ha comportato il ridimensionamento di tutte le immagini di ingresso a 1024x1024 pixel. Inoltre, attraverso una fase di Data Augmentation, sono state applicate distorsioni fotometriche, capovolgimenti e normalizzazioni per aumentare la varietà dei dati durante l'addestramento. La rete è stata addestrata end-to-end utilizzando una perdita di Dice per 100 epoche con una dimensione del batch di 4. Per l'inizializzazione, sono stati utilizzati pesi pre-addestrati e l'addestramento è stato sottoposto a un fine-tuning preciso per il nostro set di dati.

Nella fase di inference, la rete precedentemente addestrata è stata impiegata per ottenere segmentazioni automatiche delle sezioni 2D, estratte dai volumi tridimensionali OCT. Queste segmentazioni automatiche sono state ottenute su singole immagini 2D. Successivamente, è stato possibile ricostruire i volumi tridimensionali a partire dalle segmentazioni ottenute. Queste maschere segmentate sono state poi sottoposte a una fase di post-elaborazione. Questo passaggio ha permesso di migliorare ulteriormente l'accuratezza delle predizioni, ottimizzando così la qualità delle informazioni estratte dai dati OCT.

3.3 Postprocessing

Nel corso della fase di post-elaborazione, sono state condotte una serie di operazioni utilizzando il software Matlab. Queste operazioni hanno portato a un miglioramento significativo delle metriche rispetto ai volumi non post-elaborati, risultando in un

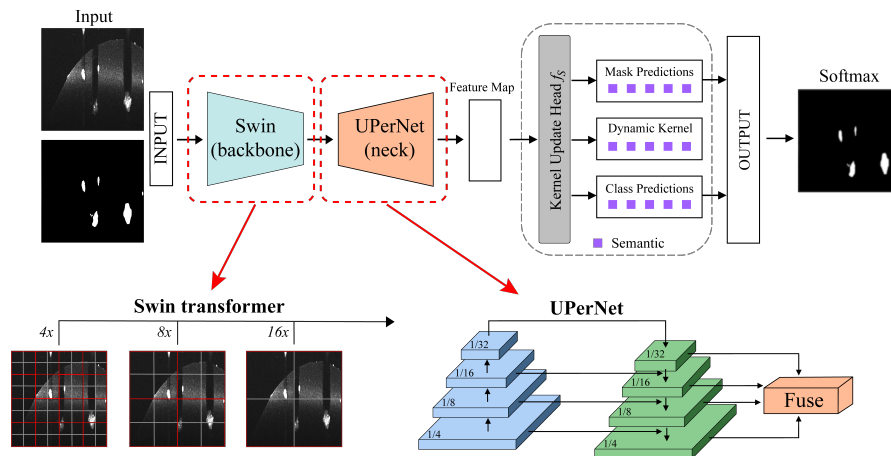


Figura 3.4: Panoramica dell'architettura K-Net utilizzata per la segmentazione degli organoidi.

incremento dell'accuratezza delle segmentazioni degli organoidi prodotti dalla rete neurale.

- **Remove Small Object (RSO):** il volume degli organoidi è stato approssimato a quello di una sfera. Successivamente, è stato calcolato un volume di riferimento, corrispondente a quello di una sfera con un diametro di $25 \mu\text{m}$. Questo valore specifico del diametro è stato fornito dal personale adetto alle acquisizioni OCT, rappresentando le dimensioni minime registrate degli organoidi nelle varie colture. Infine, sono stati rimossi dalla maschera tutti gli oggetti con un volume inferiore a quello di riferimento. Questo passaggio ha permesso di eliminare eventuali artefatti o rumori che potrebbero interferire con l'analisi successiva. La Figura 3.6 mostra il risultato ottenuto da questa fase di postprocessing.
- **Smooth3D:** in questa fase sono stati impiegati vari operatori morfologici. Questo processo consente di modificare il valore di ciascun pixel dell'immagine in base al valore degli altri pixel nel suo vicinato. La scelta della dimensione e della forma del vicinato permette di costruire un'operazione morfologica che risponde in modo specifico alle forme presenti nell'immagine in ingresso. La prima operazione eseguita è stata l'erosione, utilizzando la funzione `imerode` di Matlab e un elemento strutturale sferico di raggio $3 \times 3 \times 3$, il valore del pixel di uscita è il valore minimo di tutti i pixel del vicinato. Successivamente, è stata effettuata una dilatazione del volume con la funzione `imdilate`, utilizzando un elemento strutturale sferico di raggio $3 \times 3 \times 3$. In un'immagine binaria, il valore del pixel di uscita diventa 1 se almeno uno dei pixel nel suo vicinato ha valore 1. Infine, la funzione `imfill` è stata utilizzata per eseguire un'operazione di riempimento sui volumi. Questa operazione può essere particolarmente

utile per rimuovere gli artefatti derivanti dalla segmentazione della rete e dalle fasi precedenti di post-elaborazione. Questo approccio consente di ottenere immagini più pulite e accurate, facilitando le analisi successive. La Figura 3.7 mostra il risultato finale ottenuto da queste due fasi di postprocessing.

Attraverso la Figura 3.5, è possibile osservare i vari risultati intermedi ottenuti durante la fase di postprocessing.

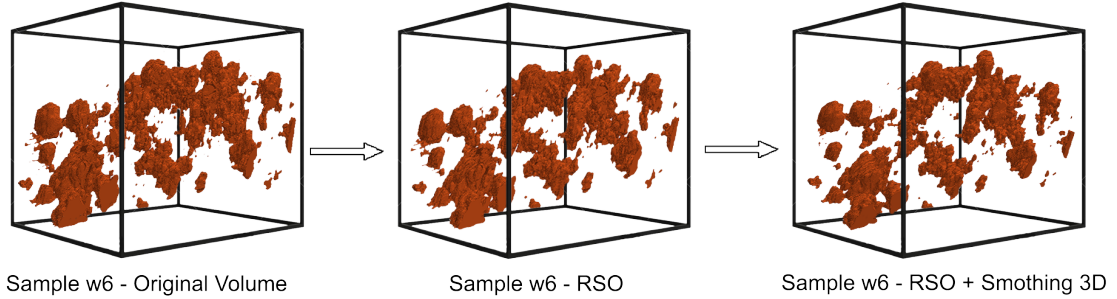


Figura 3.5: Risultati intermedi del postprocessing.

3.4 Metriche di valutazione

Nel seguente paragrafo, verranno esaminate in dettaglio le metriche di valutazione utilizzate per quantificare l'accuratezza delle segmentazioni generate dal modello di rete neurale in questione. Le metriche selezionate per questa analisi comprendono il Dice score, il Recall e la Precision. Queste sono state calcolate su immagini 2D e successivamente mediate per ottenere un valore complessivo per l'intero volume in esame. Queste metriche offrono una valutazione completa e robusta delle prestazioni del modello di segmentazione.

3.4.1 Dice score

Il Dice score, utilizzato come misura di somiglianza tra due campioni, è stato impiegato in questo studio per valutare l'adeguatezza delle previsioni del modello rispetto alla verità di riferimento. Matematicamente, il coefficiente di Dice è definito come segue:

$$Dice = \frac{2TP}{2TP + FN + FP}$$

Nel calcolo del coefficiente di Dice, i veri positivi (TP) rappresentano i pixel per i quali il modello ha correttamente predetto la classe positiva. I falsi positivi

(FP) indicano i pixel per i quali il modello ha erroneamente predetto la classe positiva, mentre i falsi negativi (FN) rappresentano i pixel per i quali il modello ha erroneamente predetto la classe negativa. Questo coefficiente tiene conto sia dell'accuratezza nella previsione della classe positiva (veri positivi) sia degli errori commessi nella previsione delle classi positive e negative. Il coefficiente di Dice si focalizza principalmente sulla regione di sovrapposizione tra la previsione del modello e la verità di riferimento. È importante notare che la dimensione della regione segmentata può influenzare i risultati complessivi del coefficiente di Dice. Pertanto, una valutazione accurata richiede un'attenta considerazione di tutti questi fattori.

La Figura 3.6 illustra le differenze tra i vari tipi di predizione.

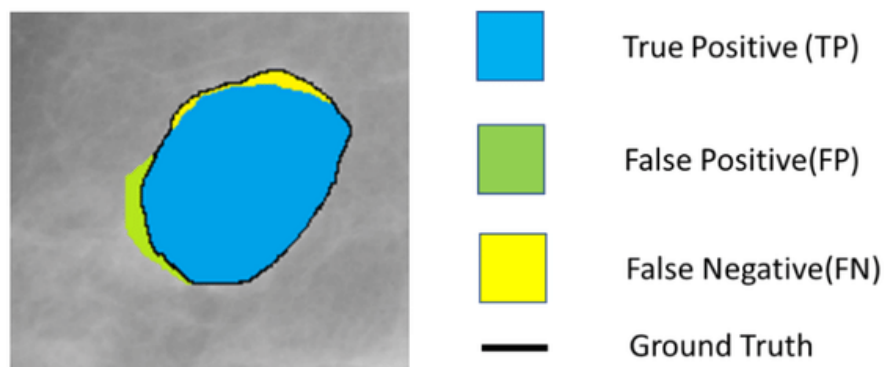


Figura 3.6: Esempio di predizione di un'immagine 2D con il relativo Ground Truth.

3.4.2 Recall

Il Recall rappresenta una metrica fondamentale nei compiti di classificazione. Questa metrica evidenzia l'abilità del modello di identificare correttamente tutte le istanze rilevanti. In termini più specifici, il Recall misura la proporzione di casi effettivamente positivi che sono stati correttamente identificati dal modello.

$$Recall = \frac{TP}{TP + FN}$$

3.4.3 Precision

La Precision rappresenta una metrica importante in contesti caratterizzati da un elevato numero di falsi positivi. Questa metrica fornisce informazioni sulla capacità del modello di identificare correttamente le istanze positive, minimizzando il rischio

di falsi allarmi. Frequentemente definita come valore predittivo positivo, la Precision quantifica la proporzione di predizioni corrette (true positive) rispetto al totale delle predizioni positive effettuate dal modello. Questa metrica è particolarmente utile per valutare l'efficacia del modello nel ridurre gli errori di classificazione.

$$Precision = \frac{TP}{TP + FP}$$

3.4.4 Valutazioni

La Tabella 3.1 presenta un insieme completo di metriche calcolate per ciascuno dei volumi nel set di test. Tre diverse configurazioni sono state considerate per il calcolo di queste metriche. In primo luogo, le metriche sono state calcolate sul volume originale, etichettato come 'BA', senza l'applicazione di alcun preprocessing o postprocessing. In secondo luogo, le metriche sono state calcolate su un volume che ha subito un preprocessing, indicato come 'PP'. Infine, le metriche sono state calcolate su un volume che ha subito sia il preprocessing che il postprocessing, compresi la rimozione di piccoli oggetti e lo smoothing 3D, indicato come 'PP+RSO+S3D'. Questa analisi fornisce una visione completa dell'effetto delle diverse fasi di processing sui volumi del set di test.

DAYS	METRICS								
	BA			PP			PP+RSO+S3D		
	DICE	RECALL	PRECISION	DICE	RECALL	PRECISION	DICE	RECALL	PRECISION
1	35.84	33.82	38.11	22.23	14.48	48.80	32.78	24.93	47.81
2	58.07	80.23	45.50	62.04	71.15	55.00	59.57	80.00	47.46
5	80.15	79.33	80.99	78.43	73.18	84.48	80.40	79.57	81.24
7	76.81	66.87	90.22	75.68	64.95	90.65	77.21	67.46	90.24
9	79.36	93.01	69.20	79.68	91.22	70.73	79.19	93.54	68.66
11	83.66	85.65	81.77	82.64	83.30	81.99	83.85	86.51	81.35
13	81.68	93.66	72.41	80.97	92.42	72.04	81.40	94.03	71.75
15	80.79	96.20	69.63	78.88	96.68	66.62	80.51	96.63	69.00
17	88.99	94.87	83.80	86.87	96.22	79.18	88.82	95.71	82.86
21	86.71	95.29	79.55	84.75	96.88	75.32	86.52	96.48	78.43

Tabella 3.1: Metriche calcolate sul test set relative al volume p1 w1. (BA) Baseline. (PP) Preprocessing. (PP+RSO+S3D) Preprocessing+postprocessing

La Tabella 3.2 riassume le metriche delle prestazioni di segmentazione per il set di test, evidenziando l'efficacia dell'algoritmo in diverse misure. Il Dice score, che si attesta a $0,730 \pm 0,217$, sottolinea l'accuratezza dell'algoritmo nei compiti di segmentazione. Le misure di Recall e Precision, rispettivamente pari a $0,788 \pm 0,257$ e $0,726 \pm 0,138$, indicano la capacità dell'algoritmo di identificare correttamente i pixel degli organoidi e di distinguerli dallo sfondo. Questi risultati dimostrano l'efficacia dell'algoritmo nella segmentazione degli organoidi.

DICE score	Recall	Precision
0.730 ± 0.217	0.788 ± 0.257	0.726 ± 0.138

Tabella 3.2: Tabella riassuntiva delle metriche di valutazione.

Nella Figura 3.7 A è rappresentato l'andamento del punteggio Dice nei diversi giorni di acquisizione. Nella Figura 3.7 B, invece, è possibile osservare il processo di segmentazione degli organoidi nei giorni 3, 9 e 17, presi come riferimento. Queste immagini permettono di verificare l'efficacia dell'algoritmo nel segmentare correttamente gli organoidi.

È importante notare che, nei primi giorni di acquisizione, le prestazioni sono notevolmente ridotte rispetto ai giorni successivi. Questo è dovuto al fatto che gli organoidi non si sono ancora sviluppati completamente e presentano dimensioni piuttosto ridotte. A causa di ciò, risultano molto simili al rumore circostante. Questo elemento contribuisce alla variabilità iniziale delle metriche, poiché l'algoritmo migliora progressivamente la sua capacità di differenziare tra gli organoidi e il rumore man mano che gli organoidi crescono.

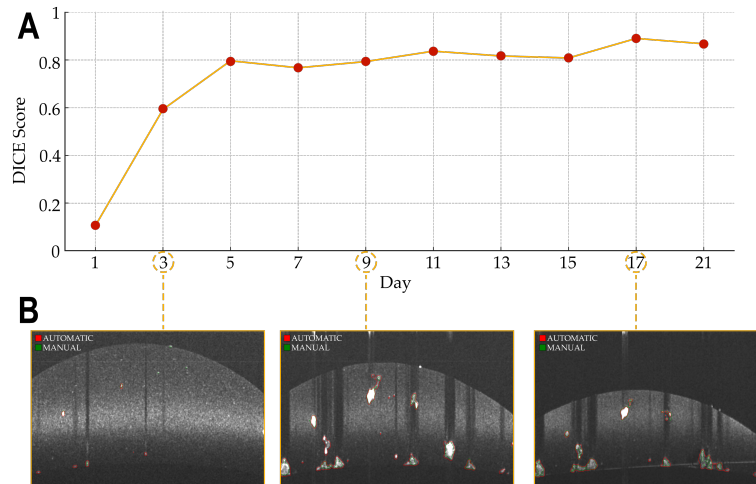


Figura 3.7: A) Andamento del Dice score nei diversi giorni di acquisizione. B) Confronto tra maschere manuali e maschere automatiche.

3.5 Algoritmo di Tracking

In questo paragrafo, viene descritto il funzionamento dell'algoritmo di tracking sviluppato per monitorare la crescita degli organoidi nei diversi giorni di acquisizione. L'algoritmo opera a partire dalle maschere predette dalla rete neurale

convoluzionale.

Il codice è stato implementato utilizzando il software Matlab 2023a. Per il corretto funzionamento del codice, è necessario disporre dei volumi tridimensionali, ottenuti ricostruendo le slice con le maschere predette dalla rete. È possibile utilizzare il codice anche direttamente con le slice dei volumi, purché siano state salvate in apposite cartelle. Nella prima parte del codice, è necessario definire alcuni parametri importanti per il corretto tracciamento degli organoidi.

3.5.1 Definizione parametri

Per il corretto funzionamento dell’algoritmo di tracking è necessario definire alcuni parametri fondamentali. In questo studio, affinché gli organoidi vengano processati correttamente, sono stati utilizzati i parametri del sistema di acquisizione. Come si può notare dalla Tabella 3.3, sono stati elencati tutti i settaggi utilizzati nell’algoritmo:

Options	
'fileType'	.mat/.png
'minDiameter'	25
'pixelSizeXY'	2.52
'pixelSizeZ'	1.26

Tabella 3.3: Parametri utilizzati per il tracking

- **'fileType'**: l’algoritmo è stato sviluppato per lavorare con due formati di file distinti. In primo luogo, tutti i volumi analizzati sono stati archiviati in file .mat. Questi file contengono sia il volume nel formato uint16, utilizzato per la realizzazione delle Maximum Intensity Projections (MIPs), sia il volume contenente le segmentazioni binarie predette, ottenute come output dalla rete neurale. Inoltre, l’algoritmo offre la possibilità di lavorare direttamente con le slices dei volumi tridimensionali. Per farlo, è necessario estrarre tutte le immagini dai volumi e salvarle in apposite cartelle. Una volta caricate queste immagini, l’algoritmo è in grado di ricostruire i volumi in 3D.
- **'minDiameter'**: questo parametro rappresenta il valore minimo del diametro utilizzato per la funzione Remove Small Objects (RSO). In questa fase, ogni organoide viene approssimato a una sfera, permettendo l’estrazione del valore del suo diametro. Se il diametro di un organoide risulta essere inferiore al 'minDiameter', l’organoide non viene considerato nella fase di analisi. Questo processo si rivela particolarmente utile per eliminare eventuali errori nella fase di predizione, garantendo così una maggiore precisione dei risultati. Nel corso

di questo studio, è stato utilizzato un valore specifico di $25 \mu\text{m}$. Questo valore è stato fornito dal personale responsabile delle acquisizioni delle colture di organoidi.

- **'pixelSizeXY'** e **'pixelSizeZ'**: I parametri in questione rappresentano le dimensioni di ogni pixel nelle direzioni X, Y e Z e sono specifici per il sistema di acquisizione OCT utilizzato. Grazie a questi valori, è stato possibile convertire il volume degli organoidi, ottenuto con la funzione 'regionprops3' di Matlab, da pixel a μm . Questo ha permesso di ottenere il diametro degli organoidi nella stessa unità di misura, facilitando così la comparazione dei risultati.

3.5.2 Tracking a doppio ramo

Nelle fasi successive di questo studio, è stato impiegato il secondo dataset, distinto da quello utilizzato per l'addestramento della rete neurale. Questo dataset comprende diverse colture di organoidi. La Tabella 3.4 elenca i nuovi campioni utilizzati, indicando il numero di volumi contenuti all'interno di ciascun campione e l'etichetta che identifica il tipo di coltura. Come si può notare, i volumi presentano etichette diverse, segnalando se le relative colture sono state sottoposte a specifici trattamenti (Treated) o meno (Untreated).

Sample	Num Volumes	Label
p1_w6	10	Treated
p1_w7	10	Treated
p1_w8	10	Treated
w4	6	Untreated
w5	6	Untreated
w7	6	Untreated
w8	6	Untreated

Tabella 3.4: Dataset utilizzato per la fase di tracking.

Prima dell'inizio del processo di tracking, sono stati definiti vari percorsi per il salvataggio di diversi file e cartelle destinati all'archiviazione dei dati. Tra questi, è stato creato un file .txt con l'obiettivo di memorizzare i parametri selezionati e, in seguito, i dati ottenuti dal tracking. Tutti i risultati, inclusi i centroidi, i volumi e i pixel di ciascuno degli organoidi tracciati, sono stati salvati in tabelle specifiche. Queste tabelle sono state poi esportate e salvate in diversi file Excel con formato .xlsx.

In questa fase, è stata adottata una metodologia di tracking a doppio ramo. A partire da un giorno di riferimento, l'analisi si articola in due fasi distinte: una ascendente e una discendente.

Il ramo ascendente procede nel tempo dal volume di riferimento, tracciando la progressione dell'organoide verso i volumi successivi, con ogni volume analizzato che diventa il nuovo riferimento per la fase successiva. In contrasto, il ramo discendente conduce un'analisi inversa, tracciando accuratamente la crescita dell'organoide fino agli stadi iniziali. La Figura 3.8 illustra la logica alla base di questo approccio di tracking.

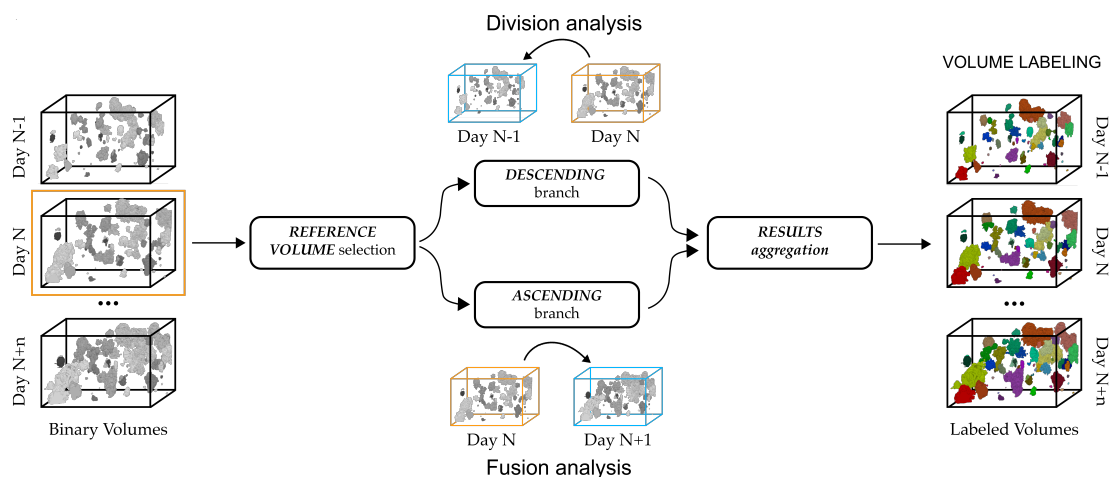


Figura 3.8: Visualizzazione dell'algoritmo di tracking.

Ogni volume è stato acquisito una volta ogni due giorni per un periodo che varia da un minimo di 13 giorni a un massimo di 21 giorni. Il giorno di riferimento è stato selezionato in base al volume centrale acquisito di ogni campione, poiché gli organoidi all'interno dei volumi risultano sufficientemente sviluppati. Questo ha permesso di adottare un approccio di tracking diverso per ciascuna fase di analisi. La Figura 3.9 illustra l'evoluzione del numero di organoidi per ogni giorno di acquisizione. È possibile notare che la maggior parte dei campioni analizzati presenta un trend decrescente per quanto riguarda il numero di organoidi presenti in ciascun volume. La scelta del giorno centrale come riferimento, ha permesso di analizzare le fusioni tra i vari organoidi nel ramo ascendente e le possibili divisioni tra di essi nel ramo discendente. Questo approccio ha fornito una visione dettagliata dell'evoluzione dinamica degli organoidi nel tempo.

Per il monitoraggio degli organoidi, è stata sviluppata una funzione denominata 'enhancedTracking'. Questa funzione accetta come input una serie di tabelle predefinite, separate per la fase ascendente e discendente. Ogni tabella contiene diverse proprietà estratte per ogni volume in analisi e i vari abbinamenti tra gli organoidi nei diversi giorni di acquisizione. Inoltre, la funzione accetta come input i parametri definiti all'inizio del processo.

Nel contesto della funzione, viene caricato il volume tridimensionale a cui viene

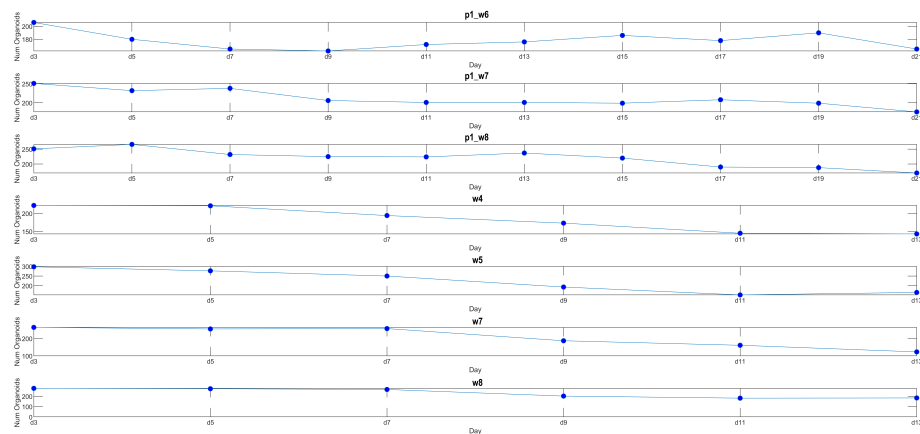


Figura 3.9: Andamento del numero degli organoidi nei diversi giorni di acquisizione.

applicato un postprocessing, precedentemente descritto. Questo postprocessing consiste in un'operazione di Remove Small Object, in cui viene utilizzato il parametro `minDiameter`, e in uno smoothing 3D. Quest'ultimo è ottenuto tramite i processi di erosione e dilatazione, seguiti dal riempimento di eventuali buchi.

Una volta processato il volume, la funzione `regionprops3` estrae le proprietà di ogni singolo organoide. Tra queste proprietà ci sono: il valore dei centroidi tridimensionali, il valore del volume, l'indice dei voxel lineari e le bounding box. Durante la prima iterazione della funzione, vengono assegnati degli ID univoci per ogni organoide. Questi ID sono fondamentali nelle fasi successive per la combinazione delle tabelle tra le fasi ascendente e discendente, nonché per il monitoraggio del loro sviluppo.

Nelle iterazioni successive, vengono estratti dalle tabelle i valori degli indici dei voxel lineari. Questi valori permettono di calcolare il Dice score tra gli organoidi di riferimento, già assegnati, e i nuovi organoidi considerati nei giorni successivi. Al termine di questa fase, si ottiene una matrice che contiene il valore del Dice score tra tutti gli organoidi di riferimento e quelli nuovi. Questa matrice risulta essere indispensabile per la valutazione delle fusioni e delle divisioni.

Di seguito, verranno dettagliatamente descritte le implementazioni delle funzioni utilizzate per analizzare i fenomeni di fusione e divisione tra i diversi organoidi nei vari istanti temporali.

Analisi Fusioni

Per la valutazione delle fusioni, è stata implementata una funzione denominata `'analyzeFusion'`. Questa funzione prende in input la matrice con i valori del Dice score e varie tabelle in cui verranno inseriti i dati del tracking.

Nella fase iniziale dell’algoritmo, viene eseguita un’operazione di Data Cleaning (come illustrato in Figura 3.10), che permette di eliminare eventuali errori di sovrapposizione. Questo processo garantisce che ogni organoide di riferimento possa essere associato a un solo organoide tra quelli nuovi ispezionati. Se un organoide di riferimento può essere associato a diversi organoidi nuovi, viene data la precedenza al nuovo organoide con volume maggiore.

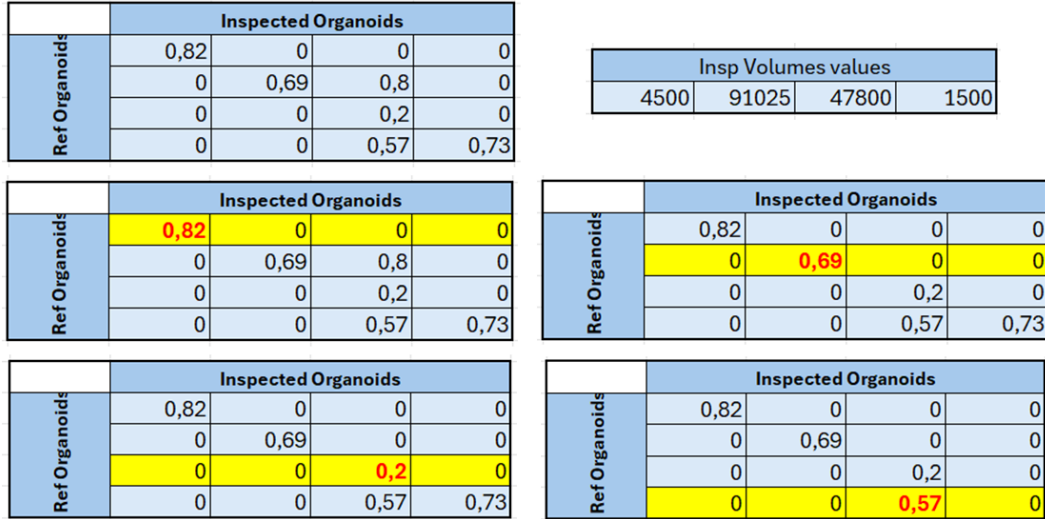


Figura 3.10: Data Cleaning della matrice con i Dice score per l’analisi delle fusioni.

Successivamente alla fase di Data Cleaning, si procede con la fase di assegnazione. In questa parte della funzione, i valori del Dice score vengono confrontati per colonna. Nel caso in cui una colonna presenti più valori, viene data precedenza all’organoide di riferimento che ha un valore di volume maggiore rispetto agli altri. Come si può osservare in Figura 3.11, l’organoide ispezionato tre è stato assegnato all’organoide di riferimento tre. Per quanto riguarda il quarto organoide di riferimento, esso non verrà assegnato. Infatti, all’interno della matrice di assegnazione, dove sono presenti tutti i match tra i diversi organoidi, verrà inserita la notazione ‘Fused_3’ per indicare che quell’organoide non verrà più tracciato in quanto è stato fuso con l’organoide tre.

Per ogni corrispondenza individuata, vengono estratte le proprietà specifiche di ciascun organoide, tra cui i centroidi tridimensionali, il volume, l’indice dei voxel lineari e le bounding box. Queste proprietà vengono successivamente salvate all’interno delle tabelle, in corrispondenza dell’ID dell’organoide assegnato.

	Inspected Organoids			
Ref Organoids	0,82	0	0	0
	0	0,69	0	0
	0	0	0,2	0
	0	0	0,57	0

Ref Volumes values			
1000	85021	43600	1200

	Inspected Organoids			
Ref Organoids	0,82	0	0	0
	0	0,69	0	0
	0	0	0,2	0
	0	0	0,57	0

	Inspected Organoids			
Ref Organoids	0,82	0	0	0
	0	0,69	0	0
	0	0	0,2	0
	0	0	0,57	0

	Inspected Organoids			
Ref Organoids	0,82	0	0	0
	0	0,69	0	0
	0	0	0,2	0
	0	0	Fused_3	0

	Inspected Organoids			
Ref Organoids	0,82	0	0	0
	0	0,69	0	0
	0	0	0,2	0
	0	0	Fused_3	0

Figura 3.11: Fase di assegnazione e analisi fusioni.

Analisi Divisioni

Per la valutazione delle divisioni, è stata sviluppata una funzione denominata 'analyzeDivision'. Questa funzione prende in input la matrice con i valori del Dice score e varie tabelle, nelle quali verranno inseriti i dati del tracking. Similmente alla funzione per la valutazione delle fusioni, anche in questo caso viene effettuata una fase preliminare di Data Cleaning, con l'obiettivo di eliminare eventuali errori nelle sovrapposizioni. Come illustrato in Figura 3.12, ogni organoide nuovo ispezionato può essere associato a un solo organoide di riferimento. Viene data la priorità agli organoidi di riferimento con un volume maggiore, mentre per gli altri verrà assegnato il valore 0. Pertanto, nella fase successiva sarà possibile assegnare lo stesso organoide di riferimento a uno o più organoidi ispezionati.

Durante la fase di assegnazione, la matrice del Dice score modificata viene analizzata per righe, valutando tutte le associazioni di un organoide di riferimento rispetto ai nuovi organoidi. Nel caso in cui un organoide di riferimento possa essere assegnato a più organoidi ispezionati, viene introdotta una nuova nomenclatura. Come illustrato in Figura 3.13, si può osservare un esempio di tracking di un organoide presente all'interno del campione w8. Per ogni organoide che ha subito fenomeni di divisione in organoidi più piccoli, l'ID viene modificato a seconda del giorno di divisione e del numero di divisioni che sono avvenute. Partendo dal giorno di riferimento d7, si nota che nel giorno d5 l'organoide iniziale si è diviso in tre organoidi e a ciascuno di essi è stato assegnato un nuovo ID univoco: 113.1, 113.2 e 113.3. Passando al giorno tre, si nota che solo l'organoide 113.1 ha subito un'ulteriore divisione e per questo il nuovo ID presenterà la nomenclatura 113.1.1 e

Ref Organoids	Inspected Organoids			
	0,5	0	0	0
	0	0,69	0	0
	0,89	0,23	0,2	0
	0	0	0,57	0,32

Ref Volumes values			
2560	31256	29850	1320

Ref Organoids	Inspected Organoids			
	0	0	0	0
	0	0,69	0	0
	0,89	0,23	0,2	0
	0	0	0,57	0,32

Ref Organoids	Inspected Organoids			
	0	0	0	0
	0	0,69	0	0
	0,89	0	0,2	0
	0	0	0,57	0,32

Ref Organoids	Inspected Organoids			
	0	0	0	0
	0	0,69	0	0
	0,89	0	0,2	0
	0	0	0,57	0,32

Ref Organoids	Inspected Organoids			
	0	0	0	0
	0	0,69	0	0
	0,89	0	0,2	0
	0	0	0	0,32

Figura 3.12: Data Cleaning della matrice con i Dice score per l'analisi delle divisioni.

113.1.2. Viene sempre data la precedenza all'organoide con il volume maggiore.

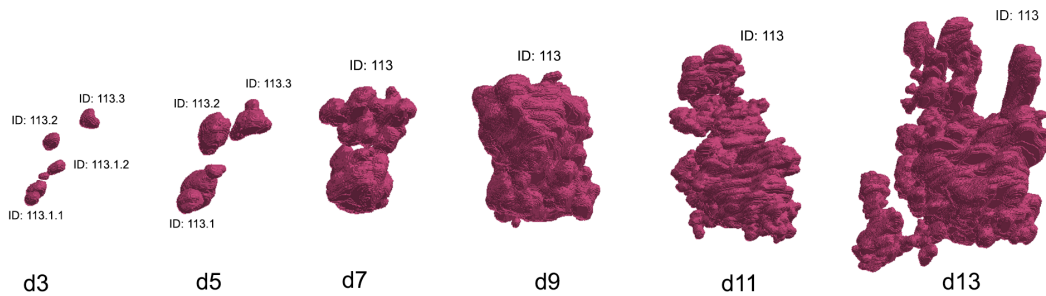


Figura 3.13: Esempio andamento fusioni/divisioni organoide 113 del sample w8.

Per ogni corrispondenza individuata, vengono estratte le proprietà specifiche di ciascun organoide, tra cui i centroidi tridimensionali, il volume, l'indice dei voxel lineari e le bounding box. Per quanto riguarda la gestione dei dati relativi agli organoidi divisi, a ciascuno di essi, oltre ai nuovi dati ottenuti, vengono assegnate tutte le proprietà ottenute prima della divisione. Questi elementi vengono successivamente salvati all'interno delle tabelle, in corrispondenza dell'ID dell'organoide assegnato.

3.5.3 Gestione Risultati

Al termine dei processi di tracking, si dispone di una serie di tabelle, ciascuna divisa in base al ramo di analisi, ascendente o discendente. Partendo dalle tabelle relative alla fase ascendente, sono stati utilizzati gli ID di questo ramo per combinare le due tabelle con i risultati, ottenendo così una singola tabella contenente tutte le proprietà di tutti gli organoidi.

Nella fase successiva, queste tabelle sono state utilizzate per l'analisi della completezza. In questa analisi, sono stati estratti tutti gli organoidi che sono stati rilevati in tutti i volumi presi in esame. Sono stati considerati anche quegli organoidi che, in un certo punto del tracking, sono stati fusi in organoidi più grandi.

Una volta ottenuti gli organoidi completamente tracciati, per ciascuno di essi, è stata valutata la coerenza dei risultati. Questa valutazione si è basata su due criteri specifici: il trend del volume e la stabilità spaziale.

Il valore di coerenza per tutti gli organoidi è stato calcolato mediando i due criteri, descritti come segue:

- **Trend del volume:** questo criterio misura la costanza con cui il volume di un organoide cresce o rimane stabile nel tempo. Si applica una regressione lineare ai dati sul volume nei vari giorni e si utilizza la pendenza della linea di regressione per calcolare un punteggio. Una pendenza positiva o nulla (che non indica alcuna diminuzione del volume) comporta un punteggio più alto, che riflette una crescita o una stabilità costante.
- **Stabilità spaziale:** questo criterio misura il grado di spostamento di un organoide rispetto alla sua posizione in un giorno di riferimento. I centroidi degli organoidi vengono confrontati con la posizione del centroide nel giorno di riferimento e le deviazioni oltre un determinato raggio di confine vengono segnalate come anomalie. Il punteggio spaziale viene calcolato in base alla percentuale di giorni in cui l'organoide rimane all'interno del raggio di confine, indicando la stabilità.

La funzione esegue un'iterazione su ciascun organoide, calcolando i punteggi di coerenza per l'andamento del volume e i punteggi di stabilità spaziale. Questi punteggi forniscono una misura quantitativa dell'evoluzione temporale e della posizione spaziale degli organoidi.

Inoltre, la funzione genera grafici che rappresentano le coordinate dei centroidi e le tendenze del volume, evidenziando eventuali valori anomali. Se l'opzione di salvataggio è abilitata, questi grafici vengono conservati come immagini per un'analisi successiva.

Il punteggio di coerenza generale viene calcolato come la media dei punteggi di

tendenza del volume e di stabilità spaziale. Questo punteggio fornisce una misura complessiva della stabilità dell'organoide.

La Figura 3.14 illustra un esempio di grafico ottenuto al termine dell'analisi di coerenza. Nella parte sinistra del grafico è rappresentato l'andamento dei centroidi; una sfera verde indica il raggio di confine oltre il quale il centroide viene considerato come outlier. Nella parte destra del grafico è rappresentato l'andamento del volume. In ciascun grafico, sono annotate informazioni quali l'ID dell'organoide, i valori di coerenza e l'indicazione se l'organoide è stato fuso con organoidi più grandi durante l'analisi.

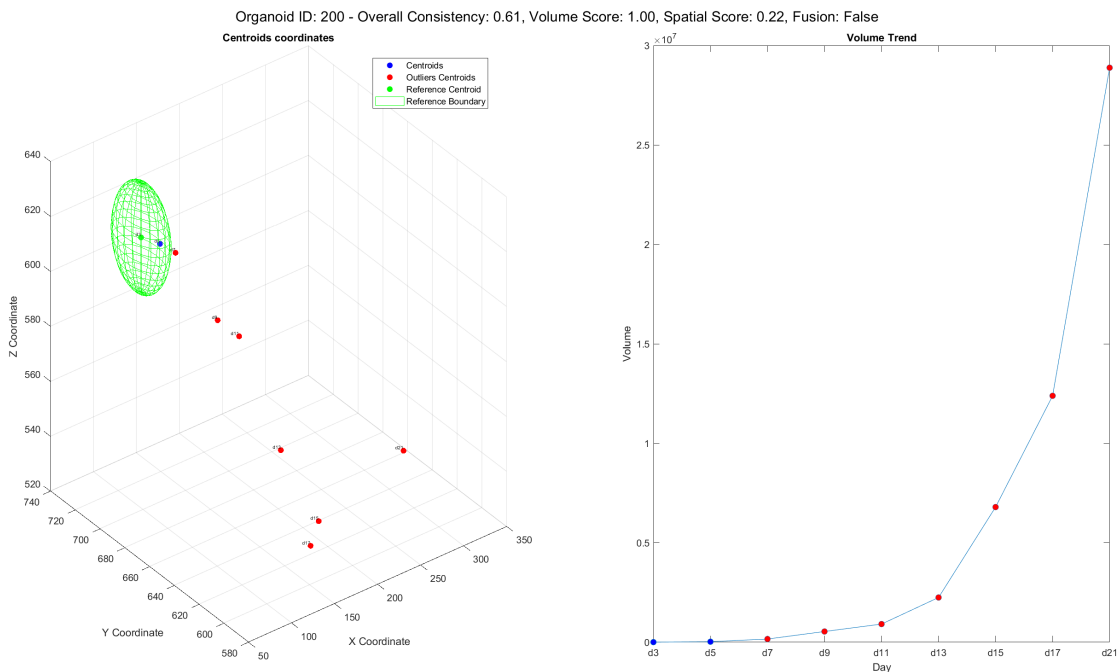


Figura 3.14: Esempio di analisi di coerenza, organoide ID:200, sample: w8

Nella fase conclusiva dell'algoritmo, tutti i dati ottenuti durante la funzione di tracking vengono salvati. Questi includono i valori dei centroidi tridimensionali e il volume degli organoidi tracciati, fusi e persi. Questi dati, contenuti nelle tabelle di Matlab, vengono esportati e salvati in file Excel nel formato .xlsx.

Vengono inoltre salvate le impostazioni definite all'inizio dell'algoritmo, tra cui: il tipo di file, il diametro minimo, le dimensioni dei pixel, il numero di volumi per ogni campione, il volume utilizzato come riferimento e il numero di organoidi completamente tracciati. Tutti questi dati vengono salvati in un file di log nel formato .txt.

Utilizzando questi dati, viene effettuata una fase di labeling. Ogni volume viene ricreato a partire dai risultati del tracking. A ogni organoide viene assegnato

un codice identificativo univoco, permettendo così di ottenere una visualizzazione tridimensionale in cui è possibile osservare l'andamento della crescita degli organoidi.

Capitolo 4

Risultati

Successivamente alla fase di segmentazione e alla generazione di maschere automatiche, l'algoritmo di tracciamento è stato applicato per delineare le dinamiche sia collettive che individuali degli organoidi. Questo processo è stato eseguito per tutti i campioni presenti nel data set.

In questo capitolo, vengono esaminati i risultati ottenuti al termine del processo di tracking. La Tabella 4.1 fornisce una panoramica generale dei dati. Viene presentato il numero totale dei volumi analizzati, il volume di riferimento impiegato come punto di partenza per il tracking, il conteggio degli organoidi che sono stati tracciati in maniera completa e la percentuale di organoidi rilevati rispetto al totale presente nel giorno di riferimento.

Sample	Label	Num of volumes	ref Volume	Tracked Organoids	% Tracking
p1_w6	Treated	10	d11	122	56.48 %
p1_w7	Treated	10	d11	173	77.23 %
p1_w8	Treated	10	d11	181	71.54 %
w4	Untreated	6	d7	100	49.50 %
w5	Untreated	6	d7	132	50.97 %
w7	Untreated	6	d7	78	30.00 %
w8	Untreated	6	d7	147	56.54 %

Tabella 4.1: Numero di organoidi completamenti tracciati per ogni volume.

4.1 Visualizzazione risultati

Dopo aver elaborato tutti i dati del tracking e averli salvati in file specifici (.xlsx e .txt), si è passati alla fase di visualizzazione 2D e 3D. Questa fase ha fornito una prova oggettiva dell'accuratezza del tracciamento di ciascun organoide. In

questa fase, utilizzando i valori dei centroidi di ciascun organoide, è stato possibile ricreare il volume tridimensionale. All'interno di questo volume sono presenti solo gli organoidi che sono stati tracciati in ogni giorno di acquisizione. A ciascun organoide è stato assegnato un ID univoco per identificarlo. La visualizzazione ha confermato l'effettiva robustezza dell'algoritmo di tracking. Attraverso la visualizzazione, è possibile osservare i diversi organoidi, ognuno con un colore diverso, che corrisponde all'ID di appartenenza. Questo permette di seguire facilmente l'evoluzione di ciascun organoide nel tempo. La Figura 4.1 offre un resoconto visivo dettagliato dello sviluppo degli organoidi in punti temporali selezionati, fornendo volumi etichettati che mappano l'andamento della crescita degli organoidi. Questa sequenza mostra visivamente un evidente aumento delle dimensioni e della complessità degli organoidi, evidenziando la capacità dell'algoritmo di adattarsi alle loro caratteristiche morfologiche in evoluzione.

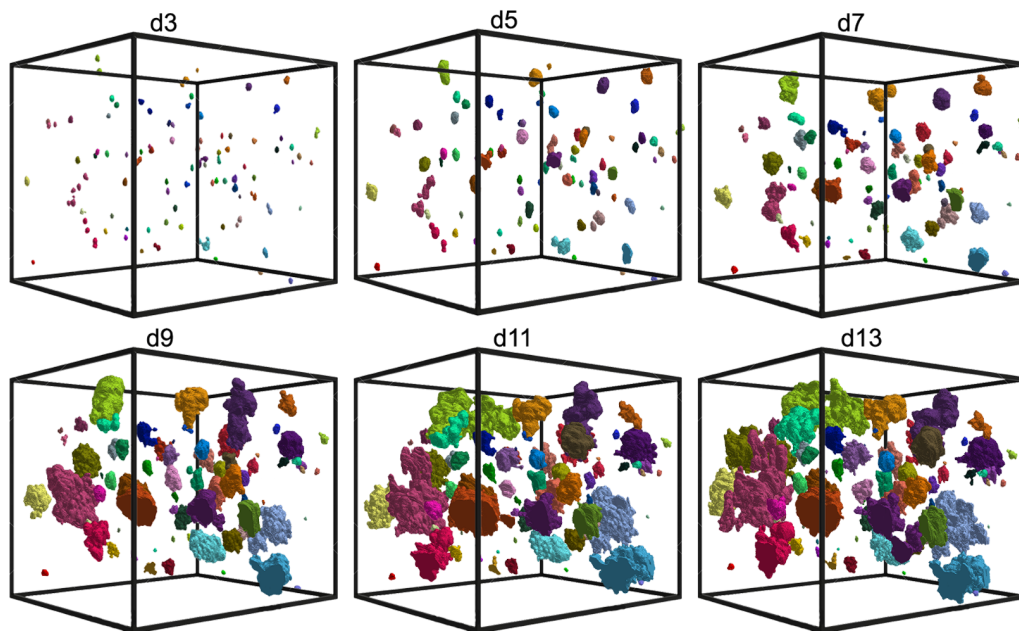


Figura 4.1: Visualizzazione 3D dei risultati del tracking ottenuti dal volume w8.

Per quanto riguarda la visualizzazione bidimensionale, si è scelto di utilizzare anche le Maximum Intensity Projections (MIPs) come alternativa alla visualizzazione tridimensionale. Nella Figura 4.2, è possibile osservare alcune delle MIPs ottenute dal volume w8 nella direzione X. L'immagine permette di osservare l'evoluzione degli organoidi e, in particolare, evidenzia i fenomeni di divisione e fusione che si sono verificati durante la crescita dell'organoide. Questo metodo di visualizzazione fornisce una rappresentazione chiara e dettagliata del processo di sviluppo degli organoidi.

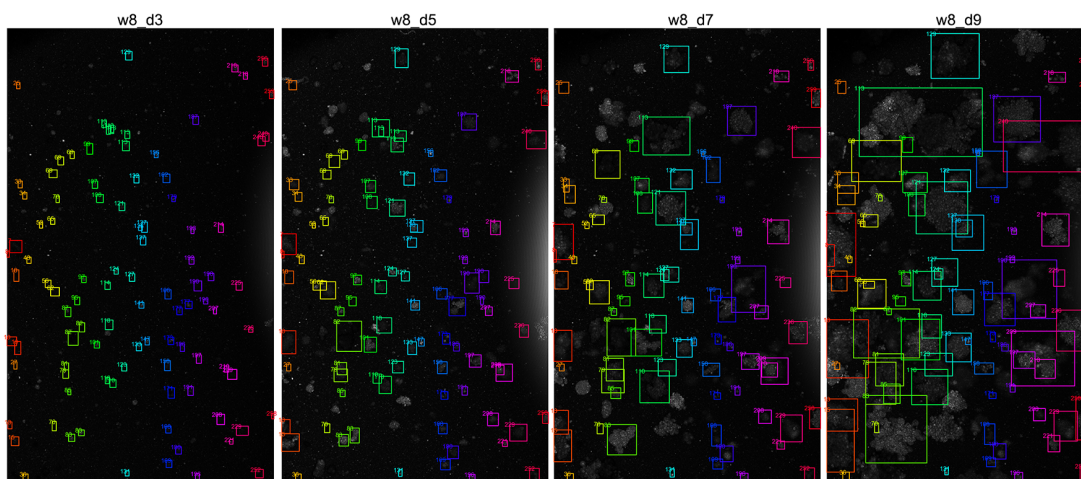


Figura 4.2: MIPs relative ai giorni d3, d5, d7 e d9 del campione w8.

4.2 Analisi di coerenza

Questa analisi è stata condotta per valutare in modo ottimale la coerenza dei dati ottenuti dal tracking, come lo spostamento dei centroidi degli organoidi assegnati durante il tracciamento e l'andamento crescente del volume nei giorni successivi allo sviluppo. I criteri di coerenza si concentrano sulla consistenza del trend di volume e sulla stabilità spaziale degli organoidi. Il primo criterio esamina la crescita o la stabilità del volume nel tempo, mentre il secondo monitora la posizione dell'organoide rispetto a un giorno di riferimento, considerando il primo giorno di acquisizione come tale. Nel processo in questione, sono stati generati grafici che permettono di osservare l'andamento dei centroidi e la crescita dei volumi. L'analisi delle fusioni ha permesso di considerare tutti gli organoidi tracciati fino a un determinato giorno e successivamente incorporati in organoidi più grandi. Questi fenomeni sono stati registrati nella tabella che contiene tutti i matching degli organoidi. Attraverso il grafico del volume, si è potuto notare un aumento significativo in determinati giorni, dovuto al fatto che un particolare organoide ha inglobato organoidi più piccoli. All'interno di tutti i grafici, è presente la notazione "Fusion: False" in caso di assenza di fusione, e "Fused + ID organoide" per indicare l'organoide in cui si è fusa. Questo permette di rintracciare l'organoide in cui si è verificata la fusione.

La Figura 4.3 illustra vari organoidi del campione w8: 104, 105, 142.1 e 142.2. Gli organoidi 142.1 e 142.2 sono stati generati attraverso un processo di divisione dall'organoide originale 142. La notazione nel titolo di ciascuna figura consente di identificare l'organoide in cui sono stati fusi. In questo caso, tutti gli organoidi menzionati precedentemente sono stati fusi nell'organoide 113. L'asse temporale

nel trend del volume indica anche il giorno in cui è avvenuta la fusione.

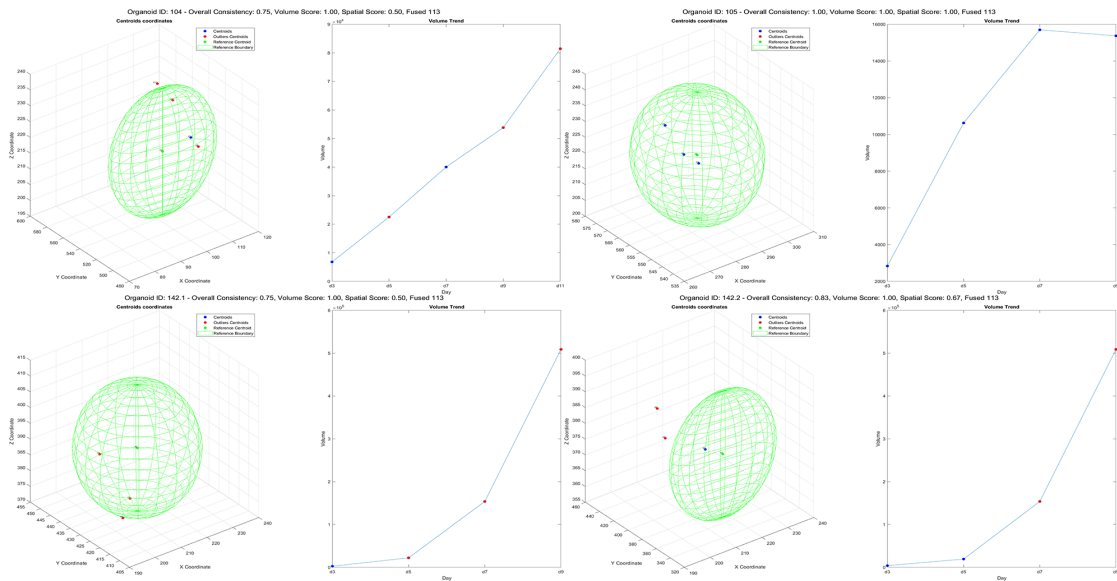


Figura 4.3: Valutazione coerenza organoidi 104, 105, 142.1, 142.2.

La Figura 4.4, riferita all'organoide 113, mostra un trend di crescita costante del volume nei giorni in cui si è verificata la fusione degli organoidi 104, 105, 142.1 e 142.2. Questo trend suggerisce un contributo significativo della fusione all'espansione volumetrica dell'organoide 113. È inoltre osservabile che esso ha subito fenomeni di divisione. Questo episodio è confermato dalla visualizzazione 3D (Figura 3.13, capitolo 3, paragrafo 3.5.2), dove si evidenzia che l'organoide 113, nei primi giorni di acquisizione, era composto da quattro organoidi differenti che si sono fusi completamente nel settimo giorno.

4.3 Crescita volumetrica

In questa fase del lavoro, sono stati analizzati tutti i risultati ottenuti dai diversi campioni del dataset. Questa fase è stata importante soprattutto per analizzare e confrontare gli i risultati ottenuti dai campioni del dataset che sono stati etichettati con 'Treated' e 'Untreated'. Le Tabelle 4.2 e 4.3 presentano, per ciascun giorno di acquisizione e per ogni campione, sia il volume totale che il volume medio registrato nello stesso giorno tra i diversi campioni. Come si può notare, a parità di giorno di sviluppo, i campioni contrassegnati come 'Untreated' mostrano un volume medio superiore rispetto ai 'Treated'. Questi risultati potrebbero essere attribuiti alla tendenza degli organoidi trattati di aggregarsi in un organoide di

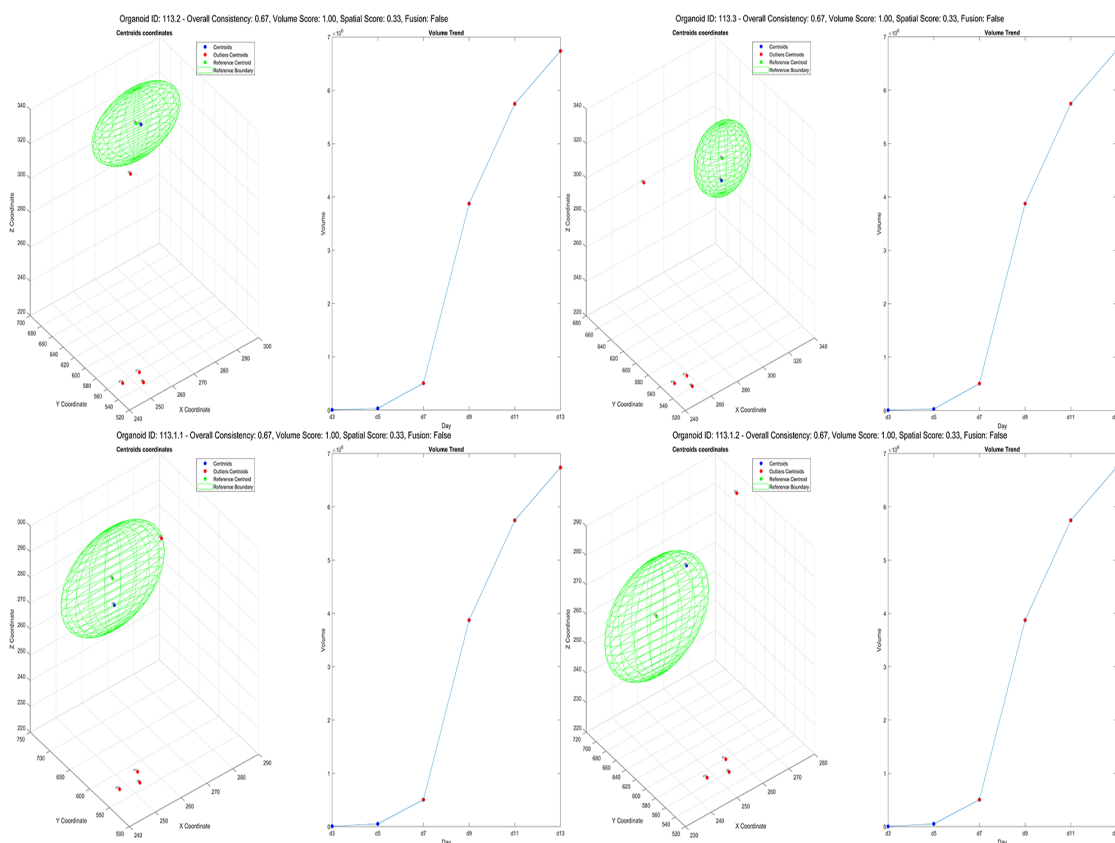


Figura 4.4: Valutazione coerenza e sviluppo dell'organoide 113.

volume considerevole. Tale fenomeno non si osserva negli organoidi non trattati, che tendono a mantenere la loro distinzione per tutta la durata del loro sviluppo nell'arco dei tredici giorni di acquisizione.

La Figura 4.5 illustra l'evoluzione della crescita dei volumi medi nei

Sample	Avg Volume Treated (μm^3)									
	d3	d5	d7	d9	d11	d13	d15	d17	d19	d21
p1_w6	4.0e+04	2.1e+05	6.6e+05	1.0e+06	1.1e+06	1.4e+06	1.8e+06	3.2e+06	4.6e+06	5.5e+06
p1_w7	3.7e+04	2.6e+05	1.0e+06	1.5e+06	2.0e+06	2.5e+06	3.8e+06	5.9e+06	9.6e+06	1.1e+07
p1_w8	3.3e+04	2.2e+05	8.6e+05	1.4e+06	1.6e+06	1.8e+06	2.4e+06	4.8e+06	9.2e+06	1.3e+07
Overall Mean	3.7e+04	2.3e+05	8.6e+05	1.3e+06	1.6e+06	1.9e+06	2.7e+06	4.6e+06	7.8e+06	1.0e+07

Tabella 4.2: Volume medio organoidi campioni "Treated"

diversi campioni, permettendo così un confronto diretto tra la crescita media dei due diversi tipi di campioni.

Nonostante i volumi 'Untreated' siano stati monitorati per un periodo di tempo inferiore, l'analisi dei dati di tracking ha rivelato che i volumi 'Untreated' w4, w5,

Avg Volume Untreated (μm^3)						
Sample	d3	d5	d7	d9	d11	d13
w4	3.6e+04	1.4e+05	6.8e+05	2.3e+06	4.1e+06	5.4e+06
w5	4.2e+04	1.9e+05	7.6e+05	2.6e+06	6.0e+06	6.9e+06
w7	4.9e+04	2.1e+05	1.0e+06	5.1e+06	8.7e+06	1.2e+07
w8	4.1e+04	1.8e+05	9.0e+05	3.8e+06	7.0e+06	8.1e+06
Overall Mean	4.2e+04	1.8e+05	8.4e+05	3.5e+06	6.4e+06	8.1e+06

Tabella 4.3: Volume medio organoidi campioni 'Untreated'

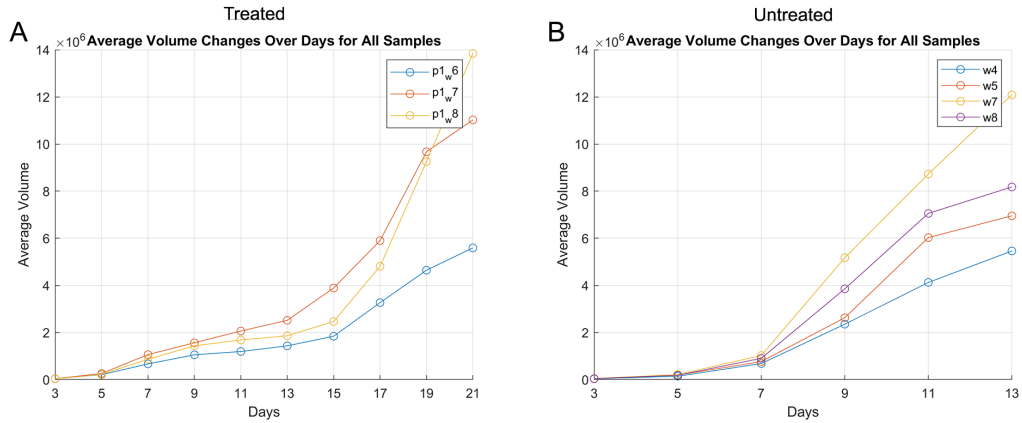


Figura 4.5: A) Trend volume medio organoidi "Treated". B) Trend volume medio organoidi 'Untreated'.

w7 e w8 mostrano un tasso di crescita superiore rispetto ai campioni 'Treated'. Il grafico evidenzia una tendenza distintiva alla crescita. In particolare, i campioni non trattati (Untreated) mostrano un tasso di crescita medio di $1.62 \times 10^6 \mu\text{m}^3/\text{giorno}$, mentre i campioni trattati (Treated) presentano un tasso di crescita medio di $1.12 \times 10^6 \mu\text{m}^3/\text{giorno}$. Questi valori rappresentano l'espansione volumetrica degli organoidi durante l'intero periodo di osservazione, fornendo una metrica cruciale per la valutazione quantitativa del loro sviluppo. Questi risultati confermano l'efficacia dell'algorithm di tracciamento nel quantificare con precisione le dinamiche di crescita.

Come precedentemente osservato, questi campioni mostrano un tasso di crescita inferiore rispetto alla controparte non trattata. Attraverso la Figura 4.6, si può notare che i campioni p1_w6, p1_w7 e p1_w8 (Treated) presentano organoidi di dimensioni considerevoli. Questa crescita esponenziale è dovuta a numerosi eventi di fusione tra organoidi avvenuti durante il loro periodo di sviluppo.

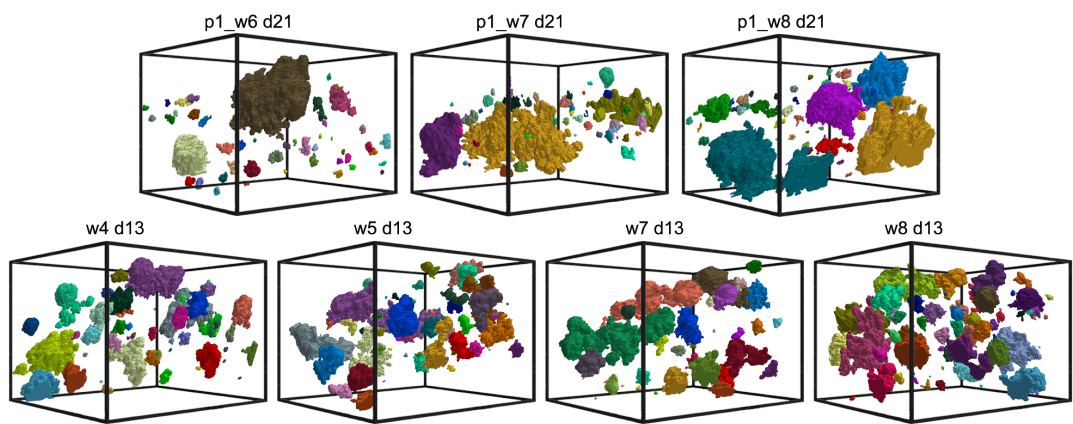


Figura 4.6: Comparazione ultimo giorno di Tracking.

Capitolo 5

Conclusione

In questo lavoro di tesi, si è potuto analizzare una serie di volumi ottenuti attraverso acquisizioni OCT di colture tridimensionali contenenti organoidi tumorali. Questi dati hanno permesso di verificare l'efficacia dell'algoritmo di tracking sviluppato, il quale ha monitorato le dinamiche sia collettive che individuali degli organoidi per l'intero periodo di acquisizione. L'analisi è stata condotta su tutti i campioni presenti nel data set, suddivisi in due categorie: 'Treated' e 'Untreated'. La categoria 'Treated' comprende colture in cui gli organoidi sono stati sottoposti a trattamento antitumorale e successivamente fatti ricrescere per valutare il fenomeno del *Drug Tolerant Persister* (DPT), tipico dei tumori, che tendono a ricrescere modificando la propria struttura del DNA per risultare più resistenti a terapie successive. I volumi 'Untreated', invece, non sono stati trattati. Ogni campione è stato segmentato utilizzando un modello di rete K-net addestrato per questo scopo, con l'obiettivo di ottenere immagini contenenti le maschere automatiche degli organoidi. A partire da esse, è stato possibile ricostruire il volume tridimensionale. L'algoritmo di tracking utilizzato adotta un approccio a doppio ramo che permette di dividere l'analisi dei volumi in due rami: ascendente e discendente. Il ramo ascendente procede nel tempo a partire dal volume di riferimento, valutando le fusioni tra organoidi sviluppati. Il ramo discendente, invece, conduce un'analisi inversa poichè traccia lo sviluppo dell'organoide fino agli stadi iniziali, permettendo di monitorare eventuali fenomeni di divisione avvenuti partendo dal giorno di riferimento. Usando i volumi segmentati, è stato possibile estrarre le proprietà principali di ogni organoide e, attraverso il Dice Score, tracciare l'evoluzione della crescita degli organoidi lungo la serie temporale, valutando la sovrapposizione tra di essi. Infine, dopo aver ottenuto tutti i matching degli organoidi nei diversi volumi, i dati sono stati salvati all'interno di specifiche tabelle. Questi risultati sono stati poi adoperati per valutare la robustezza dell'analisi, misurando il grado di coerenza dei dati stessi. Questa misura è stata effettuata analizzando il trend del valore dei volumi dei singoli organoidi e valutando la stabilità della posizione dei centroidi nei

diversi giorni di acquisizione evidenziando eventuali fusioni tra organoidi differenti. Al termine di queste analisi approfondite, è emersa una distinzione significativa tra i campioni ‘Treated’ e ‘Untreated’. I campioni ‘Untreated’, fino all’ultimo giorno di acquisizione, presentano organoidi ben differenziati con un tasso di crescita di circa $1.62 \times 10^6 \mu\text{m}^3/\text{giorno}$. Al contrario, i campioni ‘Treated’ mostrano un tasso di crescita medio di $1.12 \times 10^6 \mu\text{m}^3/\text{giorno}$. In conclusione, i campioni classificati come ‘Treated’ sono stati utilizzati per studiare il meccanismo del Drug Tolerant Persister Cancer, ovvero la capacità delle cellule tumorali di adattarsi dopo essere state sottoposte a trattamenti antitumorali. Come osservato in precedenza, questi campioni mostrano un tasso di crescita inferiore rispetto alla controparte non trattata. Tale fenomeno potrebbe essere correlato alla risposta delle cellule tumorali al trattamento antitumorale ricevuto, suggerendo che gli organoidi potrebbero privilegiare una maggiore resistenza terapeutica a discapito di una crescita più lenta.

Sviluppi futuri

Nel contesto del deep learning e delle reti neurali, una delle sfide più comuni è la dipendenza dai dati. L’addestramento di qualsiasi modello di rete neurale richiede l’accesso a grandi data set che possano caratterizzare in modo efficace il problema in esame. Uno degli sviluppi futuri per questo tipo di studio, come discusso in questa tesi, potrebbe essere l’utilizzo di nuovi set di dati in grado di migliorare la capacità, da parte della rete, di generalizzare. Inoltre, potrebbe essere utile esplorare l’uso di nuove tecniche di imaging in grado di catturare con maggiore precisione le caratteristiche morfologiche degli organoidi. Questo potrebbe consentire di ottenere dati più accurati e dettagliati, migliorando ulteriormente la qualità dell’addestramento della rete neurale. Infine, potrebbe essere interessante sperimentare nuove architetture di reti neurali. Le innovazioni in questo campo potrebbero portare a modelli più efficaci e efficienti, in grado di fornire risultati ancora più accurati aprendo nuove strade per la ricerca medica.

Bibliografia

- [1] Elisa Suarez-Martinez e Irene Suazo-Sanchez et al. «3D and organoid culture in research: physiology, hereditary genetic diseases and cancer». In: (2022) (cit. alle pp. 2, 3).
- [2] *Organoids. Biologic research and disease modeling recapitulate complexity of real tissues*. <https://www.moleculardevices.com/applications/3d-cell-models/organoids> (cit. a p. 3).
- [3] *2D Versus 3D Cell Cultures: Advantages and Disadvantages*. <https://www.mimetas.com/en/blogs/315/2d-versus-3d-cell-cultures-advantages-and-disadvantages.html> (cit. a p. 3).
- [4] Keyi Fei e Jinze Zhang et al. «Present Application and Perspectives of Organoid Imaging Technology». In: (2022) (cit. alle pp. 5–7).
- [5] Abigail J. Deloria e Sandra Haider et al. «Ultra-high-resolution 3D optical coherence tomography reveals inner structures of human placenta-derived trophoblast organoids». In: (2022) (cit. a p. 7).
- [6] *OCT Tutorial*. <https://wasatchphotonics.com/oct-tutorial/> (cit. alle pp. 8, 9).
- [7] Priyanka Malhotra e Sheifali Gupta et al. «Deep Neural Networks for Medical Image Segmentation». In: (2022) (cit. a p. 11).
- [8] *Exploring Convolutional Neural Networks: Architecture, Steps, Use Cases, and Pros and Cons*. <https://medium.com/@tam.tamanna18/exploring-convolutional-neural-networks-architecture-steps-use-cases-and-pros-and-cons-b0d3b7d46c71> (cit. alle pp. 12, 13).
- [9] Lin Shi Xiaofei Sun e Hongpeng Li et al. Yishan Luo Wei Yang. «Histogram-based normalization technique on human brain magnetic resonance images from different acquisitions». In: (2015) (cit. a p. 17).
- [10] Yan Hu e Jianfeng Ren et al. «Noise reduction by adaptive-SIN filtering for retinal OCT images». In: (2021) (cit. alle pp. 18, 19).
- [11] Zhang e Wenwei et al. «K-net: Towards unified image segmentation». In: (2021) (cit. a p. 19).

- [12] Liu e Ze et al. «Swin transformer: Hierarchical vision transformer using shifted windows». In: (2021) (cit. a p. 20).

