

POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Biomedica



Tesi di laurea magistrale

Approccio generativo tramite contenuto semantico per la simulazione di immagini istopatologiche di tessuto prostatico

RELATORI

Prof. Massimo Salvi

Ing. Francesco Marzola

CANDIDATO

Martina Del Toro

Marzo 2024

*Ai miei nonni, che sono sempre stati la mia forza.
A mia madre, che mi ha sempre spronata a continuare
anche quando volevo mollare.*

Sommario

Il carcinoma prostatico è una condizione tumorale, che colpisce gli uomini con un aumento dell'incidenza negli ultimi anni. La diagnostica di tale patologia comprende l'urinocultura, il dosaggio del PSA e l'esplorazione rettale. In relazione ad un elevato valore di PSA ed un ingrossamento della prostata, si effettua un'esame istologico del tessuto stesso per valutare la condizione tissutale. Inoltre, la diagnosi comporta l'assegnazione di un valore, detto "Gleason score", che permette di quantificare lo stato di avanzamento tumorale. Nella valutazione, però, ci sono differenti problematiche che sono tutt'ora delle open challenge: il primo è quello relativo alla variabilità inter- e intra-operatore che può modificare la valutazione che si ottiene. Per tale ragione, una delle soluzioni attualmente supportate per la semplificazione della diagnosi è quella di digitalizzare i vetrini, ottenendo quelle che vengono definite come WSI (whole slide images); il secondo è quello relativo alla mancanza di un numero sufficiente di vetrini, che permetta di costruire un dataset consistente per essere utilizzato anche con scopo educativo. Per tentare di risolvere tale problema, si è pensato di utilizzare questo lavoro di tesi, il quale ha come scopo quello di creare un algoritmo capace di sintetizzare nuove immagini istologiche, partendo da immagini reali. Per far questo, il dataset utilizzato era composto da 331 WSI, con colorazione H&E, provenienti dal Vancouver Prostate Center con le relative annotazioni definite da patologi esperti. L'idea alla base è quella di poter ottenere nuove immagini sintetiche che sia quanto più riconducibili a quelle reali, partendo dal contenuto semantico delle stesse. Per far questo è stato utilizzato un algoritmo, quale SPADE (SPatially Adaptative DENormalization), che ha come intento quello di generare delle immagini che siano semanticamente coerenti con quelle in input, partendo da quella che viene definita come maschera. Quest'ultima è una rappresentazione delle varie label che possono essere analizzate nella WSI: i gradi tumorali (condizione benigna, Gleason 3, Gleason 4 e Gleason 5), lo sfondo e tutto ciò che non era marcato. In seguito all'ottenimento delle nuove immagini sintetizzate per poter valutare la loro coerenza semantica, sono state confrontate con le immagini reali attraverso l'analisi di metriche no-reference (BRISQUE e NIQE) e con una conseguente analisi tissutale (features del primo e del secondo ordine), in particolare l'analisi è stata effettuata andando a valutare come il numero di epoche

possa condizionare la generazione di nuove immagini. Concludendo, possiamo valutare come questa sia una possibile soluzione per migliorare la robustezza e il bilanciamento di un eventuale dataset e al contempo potrebbe essere utilizzato per aumentare gli elementi utilizzabili in ambito educativo. Allo stesso tempo si potrebbe pensare di utilizzare i risultati ottenuti dall'algoritmo descritto, come elemento per l'allenamento di un discriminatore, che permetta di effettuare una diagnosi primaria.

Indice

Elenco delle tabelle	VII
Elenco delle figure	IX
1 Introduzione	1
1.1 Introduzione all'istologia	1
1.2 Anatomia e fisiologia della prostata	3
1.3 Iperplasia prostatica benigna	5
1.4 Tumore prostatico	7
1.4.1 Mezzi diagnostici	7
1.4.2 Classificazione dello stato tumorale	11
1.4.3 Fattore di rischio	13
1.5 Scopo del lavoro di tesi	14
2 Stato dell'arte	15
2.1 Whole Slide Images	15
2.1.1 Pre- e Post-processing	16
2.1.2 Stain Normalization	19
2.1.3 Applicazioni delle WSI	24
2.2 Introduzione al Machine Learning	25
2.3 Generative adversarial network (GAN)	28
2.3.1 Pix2Pix GAN	31
2.3.2 SPADE	32
2.4 Applicazioni e approcci del ML in patologia	33
2.4.1 Open challenge	34
3 Metodi	36
3.1 Presentazione del Dataset	36
3.2 Pipeline	37
3.3 Patch extraction: metodi e valutazione	39
3.4 Creazione del dataset	40

3.5	Allenamento del modello	41
3.6	Data augmentation	42
3.7	Valutazione delle metriche no reference (NR)	44
3.8	Valutazione texture analysis	45
3.8.1	Descrittori del primo ordine	45
3.8.2	Descrittori del secondo ordine	46
4	Risultati	47
4.1	Metriche No–Reference	48
4.2	Texture Analysis	54
4.2.1	Features del primo ordine	54
4.2.2	Features del secondo ordine	57
5	Conclusioni	59
A	Descrittori del primo ordine	62
B	Descrittori del secondo ordine	65
	Bibliografia	69

Elenco delle tabelle

1.1	Caratteristiche delle zone definite da McNeal per la prostata umana.	4
1.2	Caratteristiche dei Grade group con relativi Gleason score. Tabella estratta da [16].	11
1.3	Caratterizzazione del sistema TNM.	13
2.1	Sintesi delle possibili applicazioni delle WSI. Tabella estratta da [30].	24
3.1	Tabella riassuntiva della numerosità delle patch per le singole label analizzata per entrambi i metodi di estrazione.	39
3.2	Tabella riassuntiva delle aree (in valore percentuale) delle patch per le singole label analizzata per entrambi i metodi di estrazione. . . .	40
3.3	Distribuzione delle classi all'interno del Training e del Test set. . . .	41
3.4	Numerosità delle patch rispetto alle varie classi successivamente al primo e al secondo data augmentation.	43
4.1	Confronto delle metriche no-reference per il primo allenamento: valutazione senza distinzione tra le label. I risultati sono mostrati come media \pm deviazione standard.	48
4.2	Confronto delle metriche no-reference per il primo allenamento: valutazione con distinzione tra le label. I risultati sono mostrati come media \pm deviazione standard.	49
4.3	Valutazione percentuale dei valori medi della numerosità delle label sul numero totale di immagini e sull'area delle singole label rispetto all'area totale della patch. I valori sono ottenuti considerando il primo allenamento effettuato.	50
4.4	Confronto delle metriche no-reference per il secondo allenamento: valutazione senza distinzione tra le label. I risultati sono mostrati come media \pm deviazione standard.	51
4.5	Confronto delle metriche no-reference per il secondo allenamento: valutazione con distinzione tra le label. I risultati sono mostrati come media \pm deviazione standard.	52

4.6	Valutazione percentuale dei valori medi della numerosità delle label sul numero totale di immagini e sull'area delle singole label rispetto all'area totale della patch. I valori sono ottenuti considerando il secondo allenamento effettuato.	53
A.1	Definizioni matematiche dei descrittori del primo ordine.	64
B.1	Definizioni matematiche dei descrittori del secondo ordine (prima parte).	67
B.2	Definizioni matematiche dei descrittori del secondo ordine (seconda parte).	68

Elenco delle figure

1.1	Rappresentazione dello schema di funzionamento di un microscopio ottico. Immagine estratta dall'articolo [1].	2
1.2	La figura mostra l'evoluzione della divisione della prostata. (A) Rappresentazione della linea temporale relativa alla nascita delle varie suddivisioni. (B) Rappresentazione della divisione in lobi di Lowsley (1912). (C) Rappresentazione della dicotomia relativa alla sensibilità ormonale di Franks (1958). (D) Rappresentazione della divisione in zone di McNeal (1968). L'immagine è estratta da [4] . .	5
1.3	Rappresentazione di: (a) zona trasversale e (b) zona sagittale di un'immagine ecografica acquisita con TRUS di un uomo di 65 anni con cancro alla prostata. In particolare viene mostrato un possibile tumore nella zona periferica prostatica posteriore destra, come regione ipoecogena abbastanza ben definita (freccie) a ridosso della capsula. Il contorno esterno lungo la capsula posteriore è liscio il che evidenzia un cancro confinato alla ghiandola. Questa tesi è stata poi confermata, attraverso esame patologico chirurgico, come un cancro confinato all'organo in stadio T2b. Immagine prelevata da [13]. . . .	9
1.4	Rappresentazione di un'immagine con RM: l'immagine è prelevata da un paziente di 45 anni con tumore alla prostata in stadio T2b. Vengono rappresentate rispettivamente la zona (a) trasversale, (b) coronale e (c) sagittale. Le immagini sono pesate T2 e dimostrano un'area a bassa intensità di segnale (freccia) nella zona periferica ghiandolare media e nell'apice postero-laterale destro. Il tumore appare confinato alla ghiandola, poiché il margine esterno è liscio su tutti i piani [13].	10

1.5	Rappresentazione dei pattern per i cinque grade group. (a) gruppo 1: le cellule sono individuali e ben formate, (b) gruppo 2: predominanza di cellule ben formate con una minore componente poco formata, (c) gruppo 3: massiva presenza di cellule non formate e solo il 5 % di cellule formate , (d) gruppo 4: solo ghiandole scarsamente formate o ghiandole prevalentemente ben formate e componenti minori prive di ghiandole (e) gruppo 5: assenza di cellule ben formate e massiva presenza di cellule necrotiche. Immagine estratta da [17]	12
2.1	L'immagine estratta da [22] mostra: (A) WSI iniziale, (B) tessuto segmentato (in verde) identificato tramite sogliatura, (C) stesso tessuto dell'immagine in (B) ma seguito da un operatore morfologico: riempimento dei buchi.	16
2.2	Variazione del colore della stessa slide ma con l'utilizzo di due differenti scanner: (A) Aperio ScanScope V1 e (B) Aperio AT2. Immagine estratta da [22]	17
2.3	Campionamento dell'immagine: (A) con il metodo della griglia regolare, (B) con l'overlap e (C) con il campionamento randomico. L'immagine è estratta da [22].	18
2.4	Esempi di utilizzo di operatori morfologici. Questi sono usati su come filtro 11x11 e seguiti da un riempimento dei buchi: (A) dilatazione, (B) erosione, (C) apertura e (D) chiusura. L'immagine è estratta da [22].	19
2.5	Valutazione della colorazione in relazione a differenti spessori del campione.	20
2.6	Rappresentazione dell'idea di trasferimento del colore definito da Reinhard. (a) Rappresentazione del quadro di Van Gogh, da cui viene "estrapolata" l'idea del colore. (B) Immagine sorgente. (C) Immagine sorgente con il trasferimento del colore. Immagine estratta da [27].	21
2.7	Rappresentazione dei passaggi (riquadri tratteggiati) utilizzati nel modello di Khan: (1) stima delle matrici di stain, (2) color deconvolution, (3) mappatura non lineare delle statistiche dei canali e (4) ricostruzione dell'immagine. Immagine estratta da [24].	22
2.8	Rappresentazione dei passaggi dell'algoritmo WSICS: (A) immagine originale con patch per l'estrazione randomica dei campioni rappresentativi per ogni classe, (B) classificazione dei pixel per ogni classe, (C) rappresentazione delle distribuzioni cromatiche e di densità della WSI utilizzata come modello, (D) rappresentazione delle distribuzioni cromatiche e di densità successivamente all'applicazione dei pesi e (E) ricostruzione dell'immagine in RGB. Immagine estratta da [28].	23

2.9	Rappresentazione del mondo dell'intelligenza artificiale e delle sue sottobranche: il machine learning e il deep learning. Immagine estratta da [32].	25
2.10	Struttura base di una rete neurale nel deep learning. Immagine estratta da [34].	27
2.11	Diagramma a blocchi di una generative adversarial network (GAN). Immagine estratta da [36].	28
2.12	Rappresentazioni dei risultati delle diverse applicazioni delle GAN: (A) output di una CycleGAN per la image-to-image translation [37]; (B) output di un algoritmo di image super-resolution reconstruction [38]; (C) output di una CycleGAN per lo style transformation [39]; (D) output di Pix2Pix e dell'algoritmo SPADE per l'immagine generation [40].	30
2.13	Rappresentazione in dettaglio dell'algoritmo Pix2Pix (immagine estratta da [42]).	31
2.14	Rappresentazione dell'architettura del generatore considerato nel modello SPADE. (Sinistra) Struttura di un blocco residuo;(Destra) Architettura del generatore con al suo interno sono rappresentati i blocchi residui specifici. L'immagine è estratta da [40].	32
2.15	Workflow utilizzato in caso di deep learning in patologia. Immagine estratta da [43].	33
3.1	Processo di pre-processing delle WSI nel precedente lavoro di tesi.	37
3.2	Rappresentazione del processo di ottimizzazione e normalizzazione delle WSI. Si definiscono rispettivamente: (A) WSI originale, (B) WSI ottimizzata e (C) WSI normalizzata.	37
3.3	Pipeline del lavoro svolto.	38
3.4	Esempio degli elementi presenti nel Training Set, dove si ha: (A) patch, (B) mappa di segmentazione e (C) mappa d'istanza.	41
3.5	Rappresentazione delle diverse modifiche effettuate nel processo di data augmentation.	43
4.1	Rappresentazione delle immagini reali e sintetizzate, con i relativi input layer, per le sei label: Background, Benign, Gleason 3, Gleason 4, Gleason 5 e Non marcato. Le immagini sono state ottenute in seguito al secondo allenamento con il numero massimo di epoche (50).	47
4.2	Il grafico mette in relazione i dati delle metriche no-reference ottenuti per il primo allenamento, senza distinzione tra i vari label.	48
4.3	Il grafico mette in relazione i dati delle metriche no-reference ottenuti per il primo allenamento, con distinzione tra i vari label. .	50

4.4	Rappresentazione grafica dei dati relativi alle NR metrics ottenuti nel secondo allenamento. In particolare: (A) andamento delle metriche senza distinzione delle label; (B) valutazione della metrica BRISQUE con distinzione tra le label; (C) valutazione della metrica NIQE con distinzione tra le label.	53
4.5	Andamento della media, della skewness, dell'energia totale e dell'uniformità per le cinque label (Background (sfondo), Benigna, Gleason 3, Gleason 4 e Gleason 5).	55
4.6	Andamento della kurtosis, della varianza e dell'entropia per le cinque label (Background (sfondo), Benigna, Gleason 3, Gleason 4 e Gleason 5).	56
4.7	Andamento dell'autocorrelazione, del contrasto, dell'entropia differenziale, della varianza della differenza (dissomiglianza), della differenza inversa (ID), della media congiunta, dell'energia congiunta e del coefficiente di correlazione massimo (MCC) per le cinque label (Background (sfondo), Benigna, Gleason 3, Gleason 4 e Gleason 5).	58

Capitolo 1

Introduzione

1.1 Introduzione all'istologia

L'istologia è una branca della biologia che permette di analizzare i tessuti, in particolare è di grande utilità per la diagnosi in caso di tessuti patologici. L'istologia si basa sull'ausilio dei preparati attraverso la microscopia a luce trasmessa.

Per ottenere i campioni istologici il protocollo utilizzato prevede i seguenti passaggi:

1. *fissazione*: questo processo consiste nel preservare la struttura del campione, ovvero prevenendo la degradazione cellulare ed il bloccaggio del metabolismo cellulare. Il processo avviene attraverso l'immersione del campione all'interno di un fissante. Il più comune è la formalina, ovvero una soluzione acquosa con una percentuale di formaldeide pari al 37 %;
2. *inclusionione*: questo passaggio permette di ottenere un campione solido, attraverso un processo di indurimento per immersione in una soluzione alcolica tale da rimuovere totalmente l'acqua dal campione. Successivamente si ha il passaggio del campione in paraffina, la quale permea all'interno del campione, facendo sì che si abbia la solidificazione del campione stesso. In questo modo si può ottenere una struttura che viene tagliata in sezioni con spessori di 10-15 μm , grazie all'ausilio di un microtomo;
3. *colorazione*: l'ultimo passaggio è quello che rende il campione, solitamente trasparente, visibile attraverso il microscopio. Il passaggio è reso possibile attraverso l'utilizzo di alcuni coloranti, dove quelli maggiormente utilizzati sono l'ematossilina e l'eosina. Nello specifico, il campione viene inserito all'interno di una soluzione acquosa di ematossilina, la cui carica elettrostatica netta positiva permette di entrare in contatto con le cariche negative del nucleo; diversamente, l'eosina, la cui carica elettrostatica netta negativa, tende ad entrare in contatto con le strutture cellulari differenti dai nuclei.

Un elemento cardine all'interno dell'analisi istologica è il microscopio. La necessità di utilizzare un microscopio, solitamente un microscopio ottico a campo chiaro, si ha a causa della limitata risoluzione con cui l'occhio umano riesce a distinguere degli elementi nel visibile, la quale è di 0,2 mm. Poiché l'istologia tratta elementi a livello micrometrico, il microscopio risulta essere un elemento fondamentale, potendo ingrandire le immagini e rendere visibile gli elementi contenuti all'interno del campione d'analisi.

Il microscopio ottico è un elemento caratterizzato da diverse parti, quali:

- la sorgente di luce;
- delle lenti che permettono di condensare la luce sul campione;
- un ripiano su cui viene poggiato il vetrino;
- uno o più obiettivi che catturano la luce che è passata attraverso il campione;
- una lente oculare per esaminare direttamente l'immagine prodotta.

Uno schema di funzionamento del microscopio ottico a campo chiaro è mostrato in figura 1.1

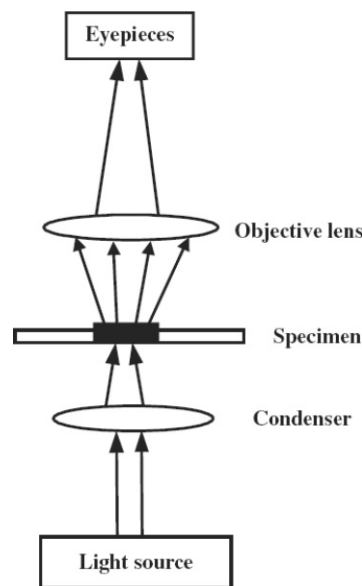


Figura 1.1: Rappresentazione dello schema di funzionamento di un microscopio ottico. Immagine estratta dall'articolo [1].

1.2 Anatomia e fisiologia della prostata

La prostata è una ghiandola inserita all'interno dell'apparato genitale maschile, nello specifico è situata all'interno della piccola pelvi maschile, con la funzione di provvedere alle proteine e agli ioni necessari alla formazione del liquido prostatico, uno dei costituenti del liquido seminale. La sua forma, in condizioni normali, è quella di una piramide con la base nella parte superiore e l'apice nella parte inferiore, rispettivamente collegate con la vescica e con la parte prostatica dell'uretra. È possibile che la forma possa variare da una forma piramidale ad una "a ciambella" nel caso di ipertrofia. Il suo massimo sviluppo si ha in fase adulta, al raggiungimento dei 25 – 30 anni di età [2], facendo sì che raggiunga un diametro di circa 4 cm, con un'altezza di 3 cm ed una profondità di 2 cm per un peso di 10-20 grammi in soggetti normali ma con possibilità di aumento nel caso di pazienti soggetti a IPB, ovvero iperplasia prostatica benigna.

L'architettura istologica della prostata è quella di una ghiandola duttale, ovvero caratterizzata dalla presenza di canali o dotti anatomici, ramificata i cui dotti sono immersi in uno stroma fibromuscolare. Quest'ultimo risulta essere presente maggiormente nella zona anteriore della prostata, quella priva di ghiandole e al contempo risulta essere un elemento fondamentale per quella che viene definita capsula. Diversamente, alla base si trovano le vescicole seminali, le quali subiscono la confluenza con i dotti deferenti su entrambi i lati, andando a formare il dotto eiaculatore.

La prima suddivisione della prostata si ha nel 1912 con gli studi Lowsley [3], il quale divide la ghiandola in cinque lobi:

- il lobo anteriore è quello che si trova anteriormente all'uretra, caratterizzato da tessuto fibromuscolare ed una scarsa quantità di tessuto ghiandolare;
- il lobo posteriore è quello che si colloca posteriormente all'uretra e inferiormente ai dotti eiaculatori, il quale risulta essere palpabile tramite esame rettale;
- i lobi laterali sono quelli che costituiscono la maggior parte della ghiandola prostatica e sono posizionati lateralmente all'uretra;
- il lobo medio, infine, è quello che si trova tra l'uretra e i dotti eiaculatori.

Successivamente, nel 1958, con Franks [4] si giunse ad una nuova divisione definita come "dicotomia della sensibilità agli ormoni sessuali". In questo caso, la prostata viene divisa in ghiandola interna e ghiandola esterna, rispettivamente considerando la prima come sede più comune per l'iperplasia prostatica mentre la seconda come sede più comune per il carcinoma prostatico. Questa suddivisione, però, non è particolarmente utilizzata in quanto non soddisfa pienamente le esigenze dell'imaging medico.

Infine, con McNeal JE [5] e il suo studio del 1981, si è giunti ad una distinzione delle zone della prostata non solo dal punto di vista istologico ma anche dal punto di vista anatomico. Nello specifico, si ha la divisione della prostata in quattro zone, le quali sono valutate partendo dall'uretra prostatica e ognuna risulta essere caratteristica da una specifica architettura. Le zone sono:

- la zona periferica (PZ), corrispondente ai lobi laterali della divisione di Lowsley, forma un disco di tessuto e costituisce circa il 70 % dell'intera ghiandola ed è la zona in cui generalmente si ha l'insorgenza di carcinomi;
- la zona centrale (CZ), corrispondente al 25 % del volume totale, è circondata dalla zona periferica nella parte distale e circonda i dotti eiaculatori. Differentemente dalla zona periferica, quest'ultima risulta essere più resistente all'insorgenza di carcinomi;
- la zona di transizione (TZ), costituente circa il 5 % del volume, avvolge l'uretra preprostatica. Questa zona è totalmente interna, avvolta anteriormente dallo stroma fibromuscolare, posteriormente è in contatto con la zona centrale e nella zona laterale e posteriore con la zona periferica;
- la zona periuretrale risulta essere una frazione della zona di transizione. Risulta essere costituita da piccoli dotti e acini non ancora sviluppati.

Una sintesi delle caratteristiche delle zone di McNeal appena descritte sono mostrate in tabella 1.1, la quale è estratta da [6].

Zona Centrale (CZ)	Volume in condizioni normali (%)	25
	Epitelio	Complesso e con larghe ghiandole poligonali
	Stroma	Compatto
	Possibilità di adenocarcinoma (%)	5
	Possibilità di iperplasia benigna (%)	–
Zona di transizione (TZ)	Volume in condizioni normali (%)	5
	Epitelio	Semplice e con piccole ghiandole rotonde
	Stroma	Compatto
	Possibilità di adenocarcinoma (%)	25
	Possibilità di iperplasia benigna (%)	100
Zona periferica (PZ)	Volume in condizioni normali (%)	70
	Epitelio	Semplice e con piccole ghiandole rotonde
	Stroma	Libero
	Possibilità di adenocarcinoma (%)	70
	Possibilità di iperplasia benigna (%)	–

Tabella 1.1: Caratteristiche delle zone definite da McNeal per la prostata umana.

Infine la rappresentazione delle zone definite da McNeal sono rappresentate in figura 1.2.

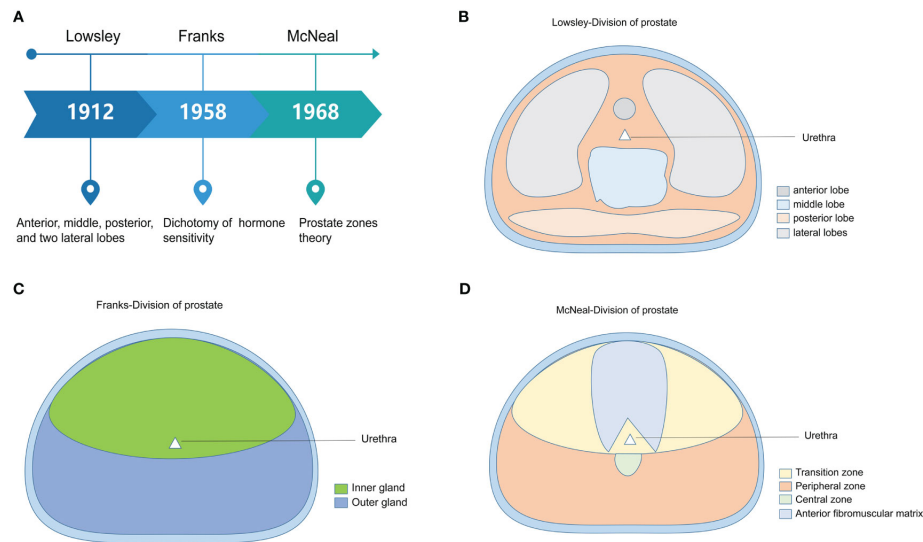


Figura 1.2: La figura mostra l'evoluzione della divisione della prostata. (A) Rappresentazione della linea temporale relativa alla nascita delle varie suddivisioni. (B) Rappresentazione della divisione in lobi di Lowsley (1912). (C) Rappresentazione della dicotomia relativa alla sensibilità ormonale di Franks (1958). (D) Rappresentazione della divisione in zone di McNeal (1968). L'immagine è estratta da [4]

1.3 Iperplasia prostatica benigna

L'iperplasia prostatica benigna (IPB) è definita come la proliferazione di cellule di stroma all'interno della ghiandola prostatica, stimolata dagli ormoni maschili, quali il testosterone e il suo derivato, il diidrotestosterone (DHT). Tale proliferazione si manifesta come un aumento in volume delle cellule dello stroma, causando un aumento in volume dell'intera ghiandola. Tale condizione genera una compressione a livello dell'uretra causando una restrizione a livello del flusso urinario. Secondo la divisione della prostata in zone, definita da McNeal, la IPB ha origine nella zona di transizione, ovvero quella che circonda l'uretra prostatica.

L'aumento di volume è una condizione fisiologica dovuta all'avanzamento dell'età e per tale motivo che questa patologia si verifica soprattutto in pazienti con età superiore ai 50 anni, anche se non mancano casi di pazienti con età nettamente inferiore (30 anni) che soffrono della stessa. L'IPB viene analizzata in seguito

alla comparsa di alcuni sintomi, che caratterizzano soprattutto il tratto uretrale inferiore, definiti come LUTS (lower urinary tract symptoms), i quali sono:

- disuria (difficoltà alla minzione);
- urgenza minzionale;
- aumento della frequenza delle inzioni diurne (pollachiuria) e notturne (nicturia);
- esitazione minzionale;
- senso di incompleto svuotamento della vescica;
- incontinenza.

La valutazione dell'iperplasia si ha attraverso l'ausilio del IPSS (International Prostate Symptom Score) [7], il quale permette di quantificare ogni sintomo rispondendo a quesiti specifici con un valore da 0 (mai) a 5 (sempre). I valori definiti vengono poi sommati tra loro, ottenendo al massimo un valore pari a 35. Nel caso in cui si ottenga un valore compreso tra 20 e 35, vengono poi effettuati i test di conferma della diagnosi.

Nella diagnosi di IPB vengono utilizzati degli esami clinici mirati, quali: l'urinocoltura, il dosaggio del PSA e l'esplorazione rettale. L'esame specifico, utilizzato per diverse patologie prostatiche è il dosaggio del PSA, o antigene specifico prostatico, il quale permette soprattutto di distinguere un caso di IPB da uno di cancro alla prostata.

Il trattamento clinico dell'IPB consiste in due macrocategorie, quella farmacologica e quella chirurgica. Nel primo caso, si ha l'utilizzo di due classi farmaceutiche principali: gli antagonisti selettivi dei recettori $\alpha - 1$ e gli inibitori della $5-\alpha$ reduttasi di tipo II. I primi hanno come scopo quello di rilassare la muscolatura liscia della prostata e del collo della vescica, in modo da migliorare il flusso urinario. Tra gli effetti collaterali si riscontrano l'ipotensione e la eiaculazione retrograda. I secondi, invece, hanno il compito di inibire la metabolizzazione del testosterone in diidrotestosterone, il quale stimola la crescita della prostata, ottenendo una riduzione in volume della prostata. La possibilità di utilizzare una cura farmacologica, viene indicata soprattutto per pazienti con una prostata di grandi dimensioni (40 g) [8].

Tra gli effetti collaterali di questa metodologia si hanno la diminuzione della libido, l'impotenza e la disfunzione eiaculatoria, ma al contempo favoriscono la riduzione del valore di PSA nel giro di sei mesi, riducendo la possibilità di cancro alla prostata. Il secondo caso, quello chirurgico, è indicato per i pazienti che non hanno migliorie con la terapia farmacologica o che hanno complicazioni legate alla malattia. Si possono definire due modalità d'azione: quella laparoscopica, in caso di pazienti con

prostate relativamente grandi (75 g) o con condizioni avverse o di prostatectomia; oppure quella in microchirurgia, caratterizzata dalla transuretrale della prostata (TURP) o la vaporizzazione laser della prostata (PVP). Negli ultimi anni, inoltre, sono state sviluppate delle possibilità di utilizzo di tecniche mininvasive, quali: termoterapia, radiofrequenza e onde ultrasoniche. Con l'utilizzo di quest'ultime il miglioramento non è pari a quello della TURP ma maggiore di quello farmacologico.

1.4 Tumore prostatico

Il tumore prostatico è uno dei più comuni tumori non cutanei che colpisce gli uomini [9], con un aumento di circa 6000 casi tra il 2017 (34800) e il 2022 (40500). Circa il 70 % dei casi insorge a livello della zona periferica della prostata, solitamente insorgendo nella zona posteriore [10]. Durante la crescita, l'adenocarcinoma, forma tumorale più diffusa (95% dei casi), può diffondersi localmente andando ad attaccare direttamente le vescicole circostanti oppure localizzarsi a distanza. Nel momento in cui viene attaccato l'apparato linfatico, aumenta la possibilità di metastasi, le quali sembrano attaccare preferenzialmente le ossa.

1.4.1 Mezzi diagnostici

I sintomi che possono presagire la presenza di un tumore prostatico sono quelli definiti come LUTS, che sono tipici della condizione dell'iperplasia prostatica benigna (IPB). Difatti, la presenza di questa sintomatologia risulta scarsamente correlata con il carcinoma prostatico; diversamente, in caso di presenza metastatica, a tali disturbi si può inserire anche un dolore a livello pelvico o lombare, che può essere considerato come un segno di presenza del carcinoma.

Tradizionalmente, il tumore alla prostata viene diagnosticato in seguito ad una esplorazione digito-rettale (DRE), un esame relativo al dosaggio di PSA e in seguito ad una biopsia tramite ultrasonografia transrettale (TRUS) [11]. Il DRE è un esame sistemico, che permette di valutare, attraverso palpazione, le dimensioni o un'eventuale irregolarità della ghiandola prostatica. Differentemente, la valutazione del dosaggio di PSA (antigene prostatico specifico) avviene con un'analisi del sangue. Uno dei problemi relativi a questa analisi è l'impossibilità di caratterizzare direttamente il tumore, in quanto è una proteina specifica di tutto l'organo e non solo della condizione tumorale. Il suo valore in condizioni normali è di 4 ng per ml di sangue e questo viene definito come soglia per la valutazione della funzione prostatica. L'uso del solo PSA, però, non permette di analizzare correttamente la benignità o la malignità della condizione, ma solo il funzionamento della ghiandola.

Per questo motivo sono state effettuate differenti considerazioni:

- il valore in normo condizione del PSA non tiene conto dell'età, la quale influisce sul suo valore. Difatti, secondo lo studio condotto da Oesterling et al. [12] si raccomandò di modificare il range secondo diverse fasce di età nello specifico le soglie erano le seguenti: 2.5 ng/mL per i 40 anni, 3.5 ng/mL per i 50 anni, 4.5 ng/mL per i 60 anni e 6.5 ng/mL per i 70 anni;
- durante le analisi vengono distinte due diverse forme di PSA, quello libero (fPSA) e quello totale (tPSA), ovvero la parte di PSA che circola nel sangue trasportata da altre proteine. Il rapporto tra queste due forme viene utilizzata per analizzare la condizione in pazienti in cui il DRE risulta normale ma il valore di PSA è elevato. Per valori percentuali molto piccoli, si hanno basse possibilità di essere in presenza di una condizione tumorale;
- la densità di PSA, valutata come rapporto tra il valore totale di PSA e il volume totale della prostata, venne introdotto come mezzo per aumentare la sensibilità e la specificità del PSA.

L'ultimo passaggio della diagnosi è quella che prevede l'utilizzo di strumentazione di imaging per l'analisi e la biopsia della prostata. L'esame più utilizzato è l'ultrasonografia transrettale (TRUS), la quale viene utilizzata per molteplici scopi tra i quali l'analisi del volume della ghiandola prostatica, utilizzata per la valutazione del PSA di densità e come guida in caso di biopsia.

L'esame prevede una fase preparatoria in cui il paziente assume antibiotici ed enema. Successivamente il soggetto assume una posizione di decubito, facendo sì che la sonda dell'ecografo entri attraverso il retto permettendo di visualizzare la prostata sia sul piano trasversale sia su quello sagittale.

Il prelievo del campione avviene con un ago calibro 18 gauge in sei differenti parti della zona periferica della prostata: apice destro e sinistro, base destra e sinistra e lato destro e sinistro della ghiandola mediale. In caso di strutture sospette, il campione viene prelevato anche in tali zone.

Questa tecnologia, però, presenta delle limitazioni dovute alla sua bassa specificità e sensibilità. Difatti, risulta avere un basso contrasto, che non permette di analizzare correttamente un tessuto tumorale da uno non tumorale e analogamente la qualità stessa dell'immagine risulta essere molto bassa. Per questo motivo, il TRUS risulta essere un esame limitato se non associato all'analisi della biopsia del campione.

Un esempio di immagine ultrasonografica con TRUS è mostrata in figura 1.3.

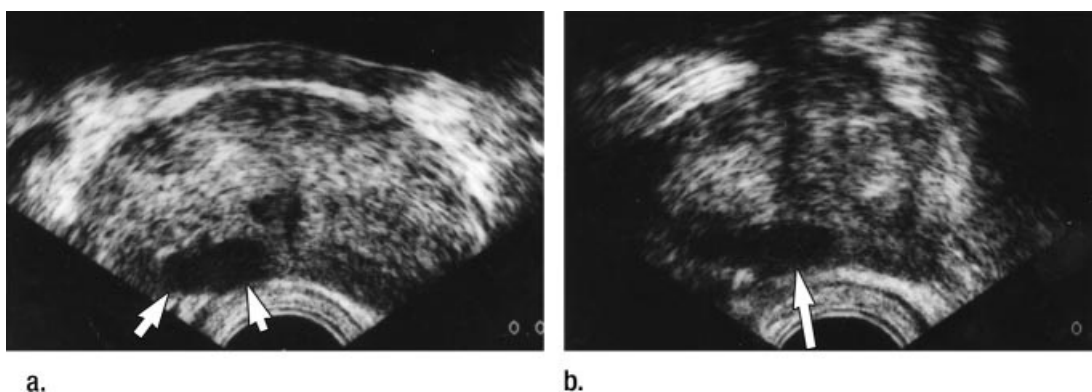


Figura 1.3: Rappresentazione di: (a) zona trasversale e (b) zona sagittale di un'immagine ecografica acquisita con TRUS di un uomo di 65 anni con cancro alla prostata. In particolare viene mostrato un possibile tumore nella zona periferica prostatica posteriore destra, come regione ipoecogena abbastanza ben definita (freccie) a ridosso della capsula. Il contorno esterno lungo la capsula posteriore è liscio il che evidenzia un cancro confinato alla ghiandola. Questa tesi è stata poi confermata, attraverso esame patologico chirurgico, come un cancro confinato all'organo in stadio T2b. Immagine prelevata da [13].

Un altro esame utilizzato per la diagnosi è l'imaging con risonanza magnetica (MRI). Quest'ultimo viene considerato nel momento in cui si ha un sospetto relativo al tumore anche in seguito ad una biopsia o ad un esame di TRUS negativo. Differentemente dall'analisi con TRUS, che permette di analizzare solo la condizione in loco, un'immagine ottenuta con RM permette di valutare anche zone distali rispetto alla ghiandola prostatica.

Nello stato attuale, l'utilizzo della RM per la valutazione del cancro alla prostata prevede l'utilizzo di un coil endorettale in unione ad un coil in zona pelvica ed un magnete a campo medio-alto (3T), in quanto per l'ottenimento di immagini con risoluzione submillimetrica, bisogna utilizzare delle immagini pesate T2. In caso di immagini pesate T1, la prostata mostra un segnale medio omogeneo non potendo così distinguere il tumore. Diversamente, nel caso di immagini pesate T2, il tumore presenta un segnale maggiore rispetto al segnale proveniente dalla zona periferica non tumorale. Poiché questa tipologia di esame permette di analizzare l'andamento distale della ghiandola prostatica, si può valutare la presenza tumorale anche a livello della zona di transizione o nella parte anteriore della zona periferica, condizione difficile in caso in cui venga utilizzato il TRUS. Questo elemento permette di andare a supportare la diagnosi, facendo sì che questo esame, seppur molto più costoso del TRUS, permetta di ottenere dei risultati migliori. Un esempio di MRI

è mostrato in figura 1.4

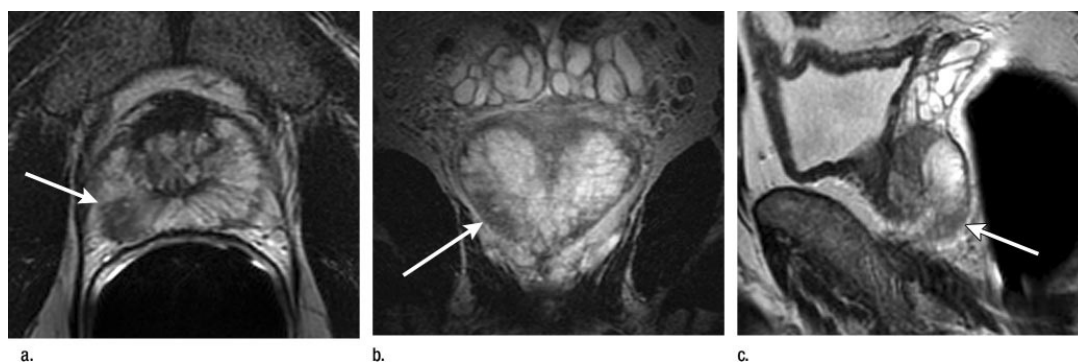


Figura 1.4: Rappresentazione di un'immagine con RM: l'immagine è prelevata da un paziente di 45 anni con tumore alla prostata in stadio T2b. Vengono rappresentate rispettivamente la zona (a) trasversale, (b) coronale e (c) sagittale. Le immagini sono pesate T2 e dimostrano un'area a bassa intensità di segnale (freccia) nella zona periferica ghiandolare media e nell'apice postero-laterale destro. Il tumore appare confinato alla ghiandola, poiché il margine esterno è liscio su tutti i piani [13].

In ultimo, possiamo considerare la MRI-TRUS, un esame che permette di combinare l'ottimizzazione nella localizzazione del tumore della MRI con la possibilità di utilizzo in tempo reale e l'ergonomia della TRUS [14]. L'idea alla base di questa fusione è quella di utilizzare le immagini ottenute con MRI come guida per la TRUS, potendo diminuire i costi e i tempi di analisi. L'idea è quella di partire da immagini di risonanza magnetica, che vengono segmentate dal radiologo e successivamente viene effettuata un'analisi ecografica in tempo reale, la quale unita alle immagini RM, permette di andare ad indirizzare correttamente la sonda per la biopsia.

Nello stato dell'arte attuale, ci sono diversi mezzi disponibili e in particolare quello che viene analizzato come esempio è quello utilizzato a Pesaro da Lacetera et al. [15], il Trinity Koelis MRI\US fusion machine.

Questo strumento prevede di esplicitare la sua funzione seguendo cinque steps:

- vengono caricate le immagini di MRI pesate T2 o DWI (diffusion weighted imaging) all'interno del dispositivo, dove vengono segmentati manualmente l'apice, la base, la ghiandola media ed eventuali punti di interesse;
- viene effettuato l'esame con TRUS, andando ad esaminare la ghiandola prostatica in tempo reale lungo i tre piani. Successivamente all'ottenimento dell'immagine viene effettuata una segmentazione come nello step precedente;
- vengono fusi i due volumi ottenuti;

- si effettua una simulazione virtuale per valutare la correttezza del volume "fuso" ottenuto;
- nel momento in cui la simulazione virtuale risulta coincidere con la zona target, viene prelevato il campione che verrà poi utilizzato per la biopsia.

1.4.2 Classificazione dello stato tumorale

In seguito al prelievo del campione viene effettuata la biopsia, la quale permette di valutare istologicamente il campione stesso. In particolare il tessuto prelevato viene colorato con coloranti quali, ad esempio, l'ematossilina ed l'eosina come nel seguente caso di studio.

I criteri maggiormente considerati per la valutazione della presenza del carcinoma sono relativi alla struttura architetturale del tessuto, nello specifico: la presenza di ghiandole troppo grandi o irregolari oppure di ghiandole piccole o cribiformi, l'assenza di cellule basali e l'allargamento dei nuclei cellulari. Per poter quantificare questi criteri vengono utilizzati due sistemi di quantificazione del grado tumorale: il Gleason score e il TNM (in inglese Tumor – Lymph node – Metastasis).

Il sistema di Gleason permette di valutare l'aggressività tumorale, assegnando il tessuto ad un gruppo compreso tra 1 e 5, ognuno dei quali assume una caratteristica ben distinta nel proprio pattern. Per assegnare correttamente ad un determinato tessuto il suo "grade group", bisogna identificare le caratteristiche tissutali delle due aree tumorali maggiormente presenti e successivamente sommarne i valori. Ad esempio, considerando un tessuto che ha nella zona maggiormente estesa un valore pari a 3 e nella zona secondaria un valore pari a 4, allora il suo punteggio finale sarà pari a 7, il che corrisponde ad un Grade group pari a 2. Le possibili classificazioni dei grade group sono mostrati in tabella 1.2.

Grade group 1	Gleason score ≤ 6
Grade group 2	Gleason score $3 + 4 = 7$
Grade group 3	Gleason score $4 + 3 = 7$
Grade group 4	Gleason score $4 + 4 = 8$ o $3 + 5 = 8$ o $5 + 3 = 8$
Grade group 5	Gleason score $= 9 - 10$

Tabella 1.2: Caratteristiche dei Grade group con relativi Gleason score. Tabella estratta da [16].

Come detto ogni gruppo ha una tipologia di pattern tissutale, visibile in figura 1.5.

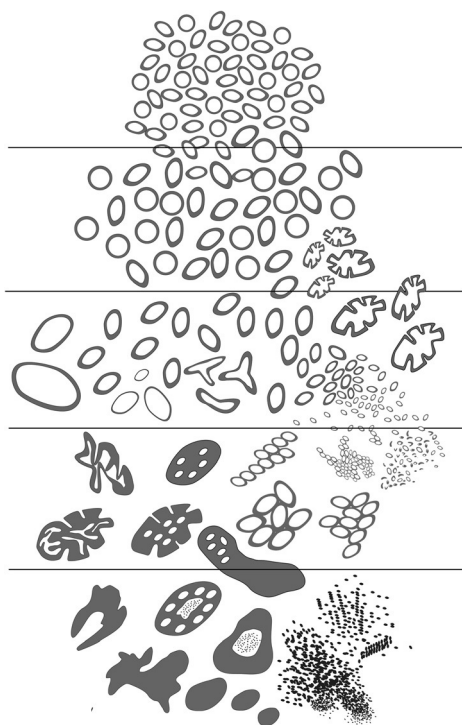


Figura 1.5: Rappresentazione dei pattern per i cinque grade group. (a) gruppo 1: le cellule sono individuali e ben formate, (b) gruppo 2: predominanza di cellule ben formate con una minore componente poco formata, (c) gruppo 3: massiva presenza di cellule non formate e solo il 5 % di cellule formate , (d) gruppo 4: solo ghiandole scarsamente formate o ghiandole prevalentemente ben formate e componenti minori prive di ghiandole (e) gruppo 5: assenza di cellule ben formate e massiva presenza di cellule necrotiche. Immagine estratta da [17]

Il sistema TNM, insieme al Gleason score e al valore di PSA, permette di determinare il livello di sviluppo del tumore. In particolare gli elementi che vengono considerati sono: il volume del tumore (T), la presenza di metastasi a livello del linfonodo (N) e la presenza di metastasi (M). Per il primo elemento viene assegnato un valore compreso tra 1 e 4, valutando il coinvolgimento della prostata e del tessuto circostante; per il secondo valore viene indicato il coinvolgimento (1) o meno (0) del linfonodo, infine, per l'ultimo valore come per il caso del linfonodo si valuta la presenza (1) o l'assenza (0) di metastasi a distanza.

Una sintesi di questa distinzione è mostrata in tabella 1.3.

Tumore primitivo (T)	TX (%)	Il tumore non può essere definito
	T0	Tumore primitivo non evidente
	T1	Tumore clinicamente non apprezzabile, non palpabile nè visibile con le immagini
	T2	Tumore limitato alla prostata
	T3	Tumore che si estende al di fuori della prostata
	T4	Tumore fisso o invade strutture adiacenti alle vescichette seminali
Metastasi ai linfonodi regionali (N)	NX	Linfonodi regionali non valutati clinicamente
	N0	No metastasi ai linfonodi regionali
	N1	Metastasi nel linfonodo regionale
Metastasi a distanza (M)	M0	Non metastasi a distanza
	M1	Metastasi a distanza

Tabella 1.3: Caratterizzazione del sistema TNM.

1.4.3 Fattore di rischio

Negli anni sono stati valutati quelli che potrebbero essere i fattori di rischio, i quali sono stati valutati andando a considerare l'incidenza del carcinoma prostatico nel tempo. Il primo elemento riscontrato è l'età. Difatti, la presenza di carcinoma si riscontra soprattutto in soggetti con età uguale o superiore ai cinquant'anni. Come detto nei precedenti paragrafi, questo elemento risulta essere correlato all'aumento del volume prostatico che si riscontra con l'avanzare dell'età. In seguito sono stati valutati altri elementi che vengono analizzati e considerati come fattori di rischio. Questi sono:

- la storia familiare: in alcuni casi di studio [18] sono stati riscontrati dei collegamenti tra il cancro alla prostata e la presenza di cancro ai polmoni. Più generalmente, si è notato come la presenza di casi di tumore all'interno della storia familiare abbia una certa influenza sul rischio tumorale;
- la razza e l'ambiente circostante: secondo alcuni studi risulta che uomini appartenenti alla razza caucasica ha una possibilità maggiore di ammalarsi rispetto ad un uomo appartenente alla razza asiatica; analogamente, un uomo di razza africana-americana ha una maggiore possibilità di ammalarsi rispetto ad un uomo di razza caucasica. Ci sono però ancora casi di studio per valutare se questo fattore di rischio sia fine a se stesso o se sia collegato all'ambiente circostante. Ad esempio, ci sono stati casi che hanno dimostrato come un uomo africano stabilito in America ha una maggiore possibilità di sviluppare un tumore alla prostata rispetto ad un uomo africano stabilito in Africa o in Asia;
- il fumo;
- una dieta ricca di latticini e priva di verdure e frutta;
- l'obesità.

1.5 Scopo del lavoro di tesi

Nello stato attuale delle informazioni relative al cancro alla prostata, un elemento importante risulta essere la valutazione del grado di aggressività dato dal gleason score e di conseguenza dalla corretta esecuzione della fase istologica, ovvero la formazione del vetrino.

Attualmente questa fase dell'analisi soffre di alcune problematiche relative alla variabilità intra- ed inter-operatore, i quali sono ulteriormente causa di artefatti. Inoltre, tale variabilità è visibile anche all'interno della diagnostica facendo sì che i gradi vengano differenziati in relazione al patologo che effettua l'analisi. Per questo motivo sono state valutate differenti possibilità, come quella relativa alla digitalizzazione dei provini e il conseguente ottenimento delle WSI (whole slide images). Queste sono alla base di nuovi studi, come quello svolto nel seguente lavoro di tesi, che hanno come scopo quello di andare a migliorare la classificazione dello stato tumorale attraverso l'utilizzo di algoritmi di intelligenza artificiale (AI) che possano generare nuove immagini partendo dal contenuto semantico dell'immagine reale.

L'idea è quella di portare un nuovo algoritmo che permetta di affiancare la figura specialistica del patologo nella classificazione tumorale, cercando di ridurre quelle che sono le problematiche precedentemente discusse. Infine, si può pensare alla possibilità di utilizzare tali algoritmi per la generazione di immagini che possano essere utilizzate con scopo didattico.

Capitolo 2

Stato dell'arte

2.1 Whole Slide Images

Le whole slide images (WSI) nascono nei primi anni duemila con lo scopo di digitalizzare i vetrini attraverso degli scanner, in modo da essere visibili a schermo con la stessa risoluzione di un microscopio ottico. Per ottimizzare l'utilizzo delle WSI, la digital pathology association (DPA) definì lo standard per ottenerle. Questo comprende i passaggi, la strumentazione e le conoscenze necessarie che risultano essere aggiuntive rispetto allo standard di ottenimento dei vetrini [19]. Generalmente i passaggi fondamentali in un processo di digital imaging sono quattro [20]:

1. cattura: acquisizione delle immagini;
2. salvataggio: archiviazione e gestione;
3. editing: manipolazione e annotazione dei dati;
4. condivisione: visualizzazione e trasmissione delle immagini.

L'ottenimento delle WSI avviene tramite scanning lineare, con lo scopo di prelevare "pezzi" di immagine, che verranno successivamente digitalizzati e assemblati per ottenere l'immagine globale. La fase di scanning è affiancata ad una fase di colorazione dei vetrini, la quale permette di classificare le WSI come: brightfield (a campo luminoso), fluorescenti o multispettrali. Con il termine WSI a campo luminoso si fa riferimento a delle immagini, le quali sono prelevate emulando l'output di un microscopio ottico ed è l'approccio più comune ed economico. Le WSI fluorescenti vengono utilizzate nel momento in cui i tessuti devono essere analizzati con la risoluzione di un microscopio a fluorescenza. Infine, nel caso delle WSI multispettrali, si tratta di una condizione di imaging in cui si prendono le informazioni attraverso lo spettro luminoso. Quest'ultima condizione può essere

applicata ad entrambe le condizioni precedenti [21].

La trasposizione digitale dei vetrini ha reso numerosi vantaggi all'analisi patologica e istologica, come ad esempio la possibilità di condivisione dei casi in esame oppure la possibilità di ridurre o potenzialmente eliminare quelli che sono i bias umani.

2.1.1 Pre- e Post-processing

In seguito all'acquisizione delle immagini si ha una fase di processing, che possiamo dividere in due fasi distinte: una prima fase di pre-processing e una seconda fase di post-processing, la quale può essere utilizzata, ad esempio, in seguito sull'output di un algoritmo di deep learning.

La fase di pre-processing segue la fase di acquisizione e ha tre differenti scopi: eliminare dati non informativi, creare un dataset consistente e consentire l'elaborazione delle immagini durante l'allenamento del modello [22]. All'interno di questa fase possiamo considerare altre sottofasi, quali: il rilevamento di artefatti o la segmentazione dei tessuti, il controllo del colore e il campionamento.

Nella fase di segmentazione e di rilevamento degli artefatti si ha come scopo principale quello di andare ad eliminare il livello informativo inutilizzabile. Per far questo si può pensare di utilizzare un processo di sogliatura che permetta di distinguere "sfondo" e "oggetto". Il metodo più utilizzato è quello di Otsu, il quale lavora sull'immagine in scala di grigi e distingue i due elementi attraverso la massimizzazione della varianza tra classi. Conseguentemente, si ottiene un processo di segmentazione, il quale può essere controllato andando ad inserire ulteriori elementi, come ad esempio dei filtri di rilevamento dei bordi. Inoltre, si possono considerare anche degli elementi morfologici che possono aiutare nel rilevamento degli artefatti. Un esempio di segmentazione e di rilevazione si ha in figura 2.1.

Il processo successivo è quello sul controllo del colore, il quale risulta essere una

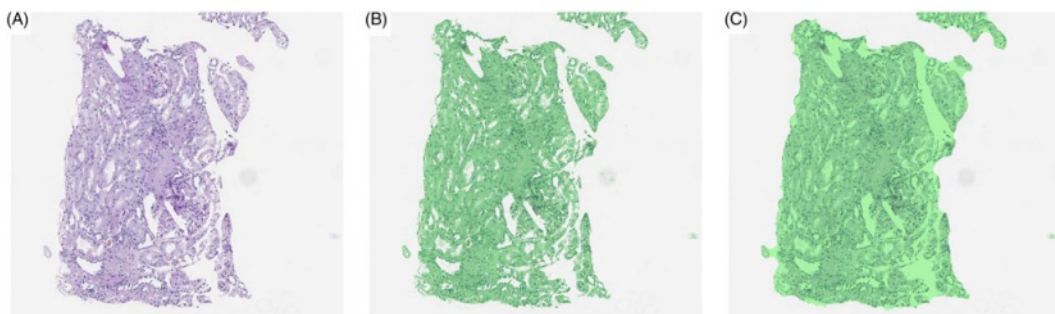


Figura 2.1: L'immagine estratta da [22] mostra: (A) WSI iniziale, (B) tessuto segmentato (in verde) identificato tramite sogliatura, (C) stesso tessuto dell'immagine in (B) ma seguito da un operatore morfologico: riempimento dei buchi.

delle open challenge attuali. Difatti, il colore tende a non essere unificato tra i vari istituti e tra i vari scanner, rendendo difficile generalizzare il modello di deep learning con i differenti dataset. Una soluzione è quella presentata da Macenko [23], il quale introdusse il principio di "color deconvolution" ovvero la separazione dell'eosina dall'ematossilina, nel caso di un tessuto con colorazione H&E.

Il processo di deconvolution di Macenko prevede la scomposizione dello spazio colore in valori singolari (singular value decomposition o SVD in inglese) per poter stabilire i vettori di colore corrispondenti a particolari coloranti. Una limitazione di questo metodo è l'impossibilità di utilizzarlo per definire un colorante che abbia tre o più componenti, in quanto non si ha più una riduzione della dimensionalità rispetto allo spazio colore. Per ovviare a questa limitazione sono stati definiti anche altri metodi, come la convoluzione del colore rispetto alla tipologia di tessuto [24]; oppure il clustering per la definizione del tessuto e in seguito la convoluzione del colore rispetto al singolo cluster [25]. Un esempio relativo al controllo del colore si ha in figura 2.2.

L'ultimo passaggio che viene incluso nella fase di pre-processing è il campiona-

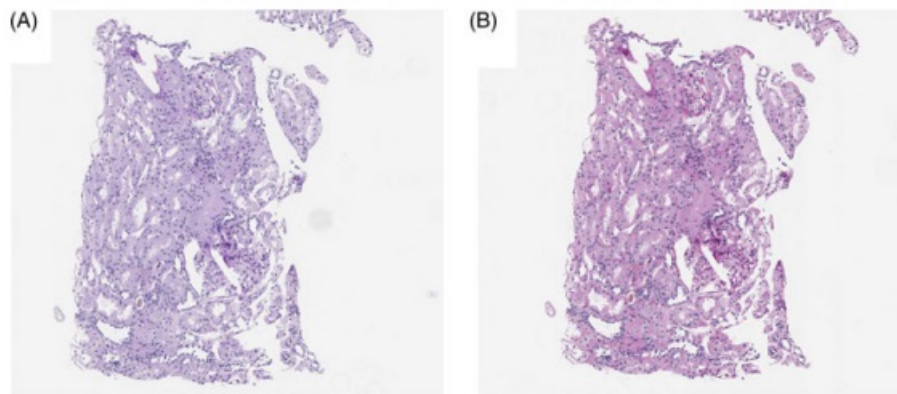


Figura 2.2: Variazione del colore della stessa slide ma con l'utilizzo di due differenti scanner: (A) Aperio ScanScope V1 e (B) Aperio AT2. Immagine estratta da [22]

mento, ovvero la divisione della slide in elementi di dimensioni idonee all'utilizzo dell'algoritmo. Questo passaggio è necessario a causa delle dimensioni delle WSI (ordine dei gigapixel), che non risultano idonee per essere utilizzate come input di un modello. La divisione dell'immagine in patch può avvenire con differenti modalità, ovvero con una griglia regolare, con sovrapposizione oppure randomicamente. Nel momento in cui si considera una divisione con griglia regolare, questa risulta essere la condizione di utilizzo più semplice ma che al contempo porta ad errori sistematici, come la rottura della struttura. Nel secondo metodo, le regioni sono prelevate con sovrapposizione con le regioni precedenti potendo così bilanciare la ridondanza degli elementi e aumentare l'efficienza computazionale. Infine, nel prelievo randomico le

patch sono prelevate in modo da coprire la regione di interesse (ROI). Un problema relativo a quest'ultimo è la possibilità di ridondanza.

Nel momento in cui la WSI viene utilizzata per allenare un algoritmo alla predizione, conviene utilizzare il metodo con overlap (sovrapposizione delle patch). In molti casi si tende ad utilizzare il metodo dello zero-padding, ovvero con l'inserimento di un numero di zeri lungo il bordo dell'immagine in modo da avere in output un'immagine che abbia le stesse dimensioni dell'immagine in input. Una rappresentazione delle tre possibili divisioni si ha in figura 2.3.

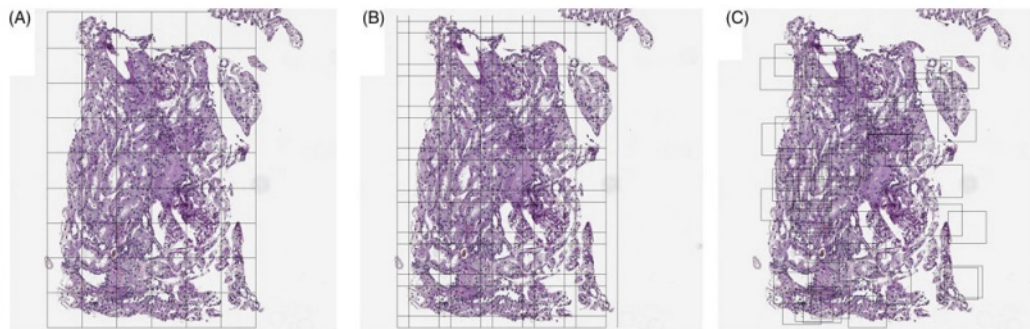


Figura 2.3: Campionamento dell'immagine: (A) con il metodo della griglia regolare, (B) con l'overlap e (C) con il campionamento randomico. L'immagine è estratta da [22].

Per quanto riguarda gli elementi di analisi del post-processing, questi vengono considerati nel momento in cui bisogna correggere dei piccoli errori che possono permettere al modello di predizione di sbagliare. Questi elementi sono:

- *riempimento dei buchi*: viene considerato per creare continuità con il bordo di un elemento, in quanto in mancanza di quest'ultimo si potrebbe avere un'errata predizione. Il ragionamento è analogo in caso in cui vengano riempiti degli elementi che sono interni al bordo;
- *utilizzo degli operatori morfologici*: gli operatori morfologici sono rispettivamente la dilatazione e l'erosione, che possono dare due differenti effetti in relazione all'ordine in cui vengono utilizzati. Nel caso in cui si utilizzi una dilatazione seguita da un'erosione, siamo in presenza di una chiusura, ovvero una rimozione dello sfondo a favore dell'oggetto (utilizzabile anche come riempimento dei buchi); invece, se si utilizza un'erosione ed in seguito una dilatazione si ha una apertura, ovvero una rimozione degli oggetti più piccoli per preservare lo sfondo;
- *rimozione degli oggetti*.

Un esempio di utilizzo degli operatori morfologici è mostrato in figura 2.4.

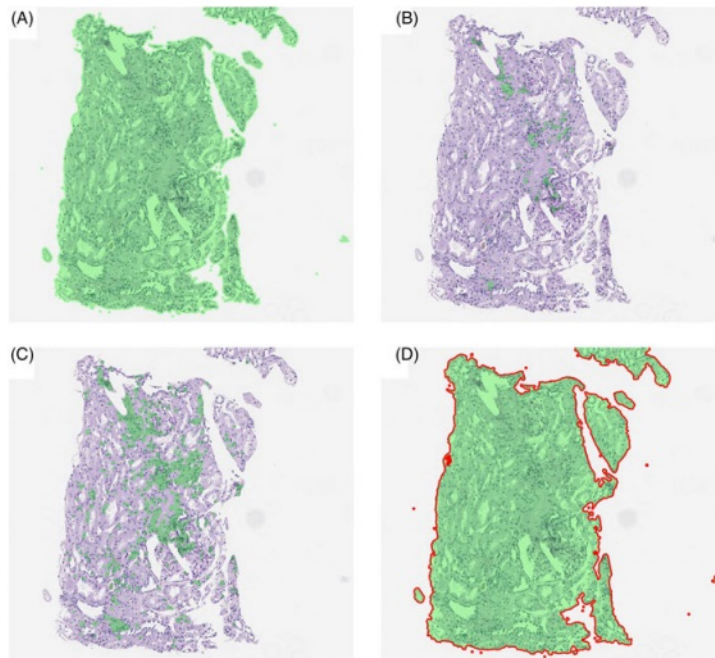


Figura 2.4: Esempi di utilizzo di operatori morfologici. Questi sono usati su come filtro 11×11 e seguiti da un riempimento dei buchi: (A) dilatazione, (B) erosione, (C) apertura e (D) chiusura. L'immagine è estratta da [22].

2.1.2 Stain Normalization

Quando si parla di WSI (Whole Slide Images), una delle limitazioni che spesso ricorre è quella relativa al colore, in particolare alla sua mancata standardizzazione e alla sua validazione nel momento in cui viene mostrata a video. La variabilità della colorazione è un elemento che compare nel momento in cui si ha la preparazione del vetrino, ad esempio a causa di una differente illuminazione o del filtraggio. Conseguentemente questa variazione si presenta anche nel momento in cui lo stesso campione viene digitalizzato. Alcune cause di variabilità della colorazione sono state valutate da Yukako Yagi nel suo articolo [26] e sono:

- lo spessore del campione: generalmente lo spessore del campione è di circa $4 - 7 \mu\text{m}$. All'aumentare dello spessore si nota come i dettagli risultino essere più scuri; differentemente per spessori minori l'immagine risulta essere molto più chiara. Un esempio è mostrato in figura 2.5;
- lo staining: questo dipende dal fatto che ogni istituto o laboratorio risulta avere delle differenti colorazioni;

- il processo di scannerizzazione;
- visualizzazione dell'immagine;
- utilizzo del display:

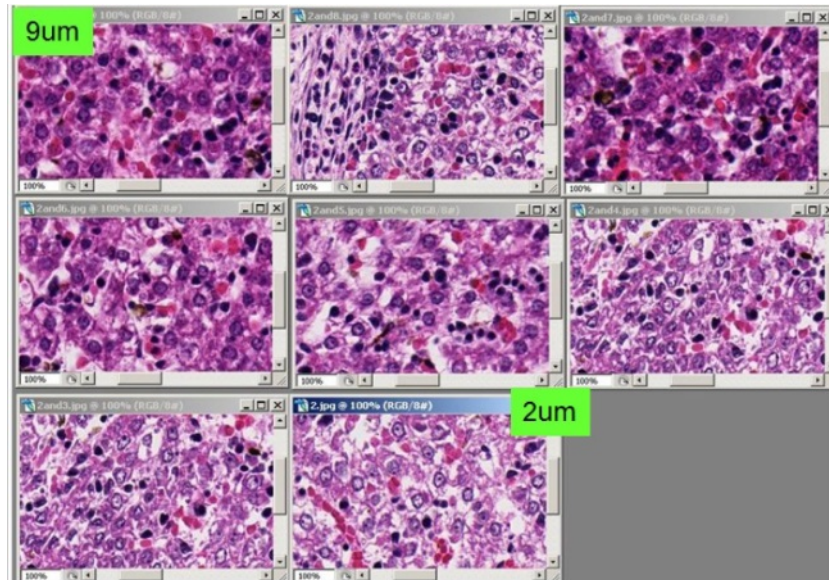


Figura 2.5: Valutazione della colorazione in relazione a differenti spessori del campione.

L'importanza della stain normalization risiede nella possibilità di aumentare l'accuratezza della fase di predizione nel momento in cui si utilizza un algoritmo di machine learning. Una possibilità, come detto nel paragrafo precedente, è l'utilizzo del metodo di Macenko. Nel tempo, però, sono state valutate differenti possibilità che prevedono l'utilizzo di algoritmi di intelligenza artificiale (AI) per la correzione del colore. Tra questi modelli ne vediamo nello specifico tre: il modello di Reinhard [27], il modello di Khan [24] ed infine il modello di Bejnordi [28].

Modello di Reinhard L'idea alla base di questo modello è quella di poter trasferire le caratteristiche cromatiche di un'immagine ad un'altra, andando a considerare che ogni pixel di un'immagine è definito dalla correlazione tra i valori dei tre canali dello spazio colore, ad esempio lo spazio RGB. In particolare, in questo studio viene considerato lo spazio colore definito da Ruderman et al. [29], il quale può essere definito come la risoluzione di un problema tridimensionale in tre problemi unidimensionali attraverso la decorrelazione tra i canali di uno spazio colore. Come nel caso di Reinhard, l'idea alla base è quella di trasferire il colore di un'immagine ad un'altra, la quale mantiene il suo contenuto semantico.

Questo viene effettuato come una serie di conversioni, che da uno spazio RGB passa ad uno spazio LMS (long,medium,short) ed infine allo spazio definito da Ruderman ($l\alpha\beta$). Un esempio dell'idea di trasferimento del colore è mostrato in figura 2.6.



Figura 2.6: Rappresentazione dell'idea di trasferimento del colore definito da Reinhard. (a) Rappresentazione del quadro di Van Gogh, da cui viene "estrapolata" l'idea del colore. (B) Immagine sorgente. (C) Immagine sorgente con il trasferimento del colore. Immagine estratta da [27].

Modello di Khan Il modello proposto da Khan prevede una mappatura non lineare regolare dei canali di colorazione, che garantisce una trasformazione dell'immagine priva di artefatti visivi. L'idea di questo modello, mostrata in figura 2.7, prevede un algoritmo costituito da quattro parti: stima della matrice di stain, color deconvolution, mappatura non lineare delle statistiche dei canali e la ricostruzione dell'immagine. La stima delle matrici di stain avviene in due fasi:

- fase di apprendimento: questa prevede l'ottenimento di uno specifico descrittore del colore (SCD) capace di quantificare la quantità di stain all'interno dell'immagine. In seguito, vengono utilizzati dei modelli di classificazione, che utilizzano contemporaneamente informazioni sui pixel (RGB) e informazioni sul colore dell'intera immagine (SCD), generando delle mappe di probabilità specifiche per ogni stain;
- fase di valutazione.

La mappatura non lineare avviene andando a calcolare per ogni canale delle immagini di input (sorgente e target) una serie di statistiche, mappando corrispondentemente quelle dell'immagine sorgente con quelle dell'immagine target attraverso una mappatura non lineare spline-based. Infine, si ha la ricostruzione dell'immagine che avviene ricombinando i canali RGB normalizzati.

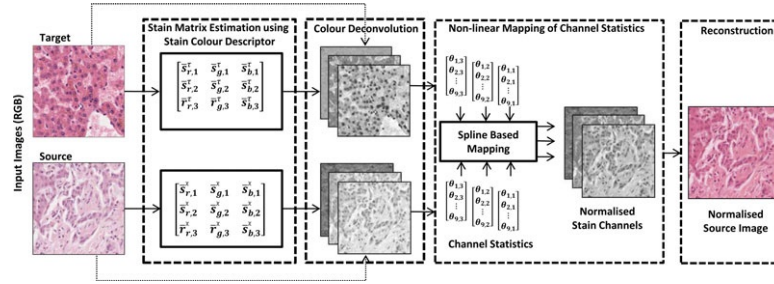


Figura 2.7: Rappresentazione dei passaggi (riquadri tratteggiati) utilizzati nel modello di Khan: (1) stima delle matrici di stain, (2) color deconvolution, (3) mappatura non lineare delle statistiche dei canali e (4) ricostruzione dell'immagine. Immagine estratta da [24].

Modello di Bejnordi Il modello di Bejnordi è un algoritmo, chiamato Whole–Slide Image Color Standardizer (WSICS), basato sulla trasformazione delle distribuzioni cromatiche e di densità per ogni singola classe nel modello di colore HSD (hue–saturation–density), ovvero tonalità–saturazione–densità. L'algoritmo prevede come input l'immagine, i cui pixel verranno classificati nelle differenti classi di stain e in seguito si hanno i seguenti sei passaggi:

- applicazione della trasformazione HSD: il modello HSD trasforma i dati RGB in dati in tre dimensioni: due cromatiche indipendenti dall'ammontare dello stain (c_x e c_y) ed una componente di densità D , linearmente dipendente dall'ammontare dello stain. Da questa trasformazione si ottiene una distribuzione $F(c_x, c_y)$ per i vari stain che compongono l'immagine. Ad esempio, per un'immagine istologica con colorazione H&E ci saranno due distribuzioni, una per l'ematossilina ed una per l'eosina;
- derivazione della distribuzione cromatica e di densità per l'ematossilina, l'esoina e lo sfondo: vengono estratti randomicamente dei campioni rappresentativi (ovvero contenenti nel 75% del tessuto) sull'ingrandimento più basso. A questi campioni viene applicata la trasformazione di Hough randomicamente ristretta, la quale permette di trovare i nuclei nella corrispondente immagine ingrandita; se il numero di nuclei trovati supera una certa soglia allora l'immagine viene classificata nelle due tinte e i pixel vengono conservati. Il processo viene effettuato fin tanto che non si giunge al voluto numero di pixel. In seguito, si avvia il processo di classificazione dei pixel nelle due componenti, in modo da ottenere le distribuzioni per ogni classe;
- trasformazione non lineare delle informazioni cromatiche: questo passaggio permette di trasformare la distribuzione cromatica di ogni classe in modo da poter combaciare con la distribuzione di un modello. Per far questo vengono

effettuati due passaggi: l'estrazione dei parametri statistici dal modello e successivamente l'allineamento con la distribuzione cromatica;

- trasformazione della distribuzione di densità delle due classi con la corrispettiva del modello;
- valutazione dei pesi del contributo dello stain sui singoli pixel, ottenendo le distribuzioni cromatiche e di densità finali;
- applicazione della trasformazione HSD inversa, ovvero il passaggio che permette di tornare ad un'immagine in RGB.

La sintesi dei passaggi di questo modello sono mostrati nella figura 2.8.

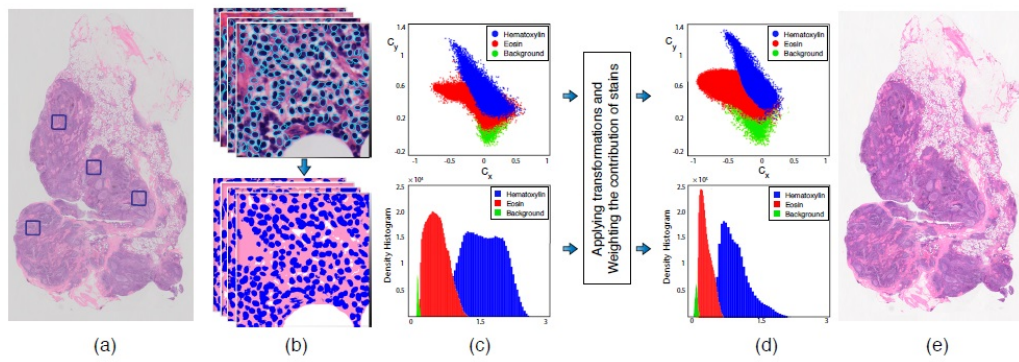


Figura 2.8: Rappresentazione dei passaggi dell'algorithmo WSICS: (A) immagine originale con patch per l'estrazione randomica dei campioni rappresentativi per ogni classe, (B) classificazione dei pixel per ogni classe, (C) rappresentazione delle distribuzioni cromatiche e di densità della WSI utilizzata come modello, (D) rappresentazione delle distribuzioni cromatiche e di densità successivamente all'applicazione dei pesi e (E) ricostruzione dell'immagine in RGB. Immagine estratta da [28].

2.1.3 Applicazioni delle WSI

Le applicazioni delle WSI possono essere racchiuse macroscopicamente in tre grandi categorie:

- applicazioni cliniche: tra queste si può considerare la diagnosi primaria, ovvero l'utilizzo delle WSI rispetto ai classici vetrini, per definire la diagnostica; analogamente, queste vengono spesso utilizzate per semplificare le consultazioni nel momento in cui si hanno dei referti precedenti alla diagnostica attuale. In questo modo si può facilitare la consultazione intra-operatore. Un'altra importante applicazione è la possibilità di lavorare sullo stesso paziente anche se si è in due locazioni differenti, quindi avere una forma di telepatologia che permetta il contatto tra i diversi medici. Questo aspetto è, inoltre, utile per il paziente il quale può conservare e avere a disposizione differenti referti allo stesso tempo. L'ultima applicazione è quella della validazione clinica, ovvero affiancare la diagnosi del patologo a quella definita da un algoritmo di intelligenza artificiale;
- applicazioni non cliniche: tra queste quella più considerata è quella di strumento accademico per l'apprendimento degli studenti potendo, inoltre, creare dei dataset agibili per gli studenti e gli insegnanti. Si può, inoltre, pensare all'utilizzo delle WSI per valutare la qualità dei vetrini;
- applicazioni di ricerca.

Una sintesi delle possibili applicazioni è mostrata in tabella 2.1.

Cliniche	Non cliniche	Ricerca
Diagnosi primaria	Presentazioni accademiche	Analisi di immagini
Consultazioni	Pubblicazioni	Microarray tissutali
Consultazioni intraoperatore	Garanzia di qualità	Multiplex staining
Microdissezione automatica	Educazione	Scoperta di biomarker
Conferenza multidisciplinare	Social Media	Ricostruzione 3D

Tabella 2.1: Sintesi delle possibili applicazioni delle WSI. Tabella estratta da [30].

2.2 Introduzione al Machine Learning

L'intelligenza artificiale è definita da Shapiro [31] come "un campo della scienza e dell'ingegneria che si occupa della comprensione computazionale di ciò che è comunemente chiamato comportamento intelligente e della creazione di artefatti che mostrano tale comportamento".

La prima comparsa dell'idea di intelligenza artificiale si ha nel 1950 con Turing, il quale definisce un comportamento intelligente quello di un computer che ha l'abilità di comportarsi come un essere umano. Egli riuscì a rappresentare questa idea attraverso quello che prende il nome di "test di Turing". Questo prevede che ci sia un soggetto C diviso da altri due soggetti A e B, di cui uno è un computer. In seguito a delle domanda a cui entrambi devono rispondere, il soggetto C deve essere in grado di non distinguere chi tra i due soggetti è il computer.

Con gli anni, l'AI ha subito numerosi sviluppi che hanno portato alla definizione di due sottoclassi, quali il machine learning (ML) e il deep learning (DL). Un'idea della divisione e delle caratteristiche generali delle tre categorie è mostrata in figura 2.9.

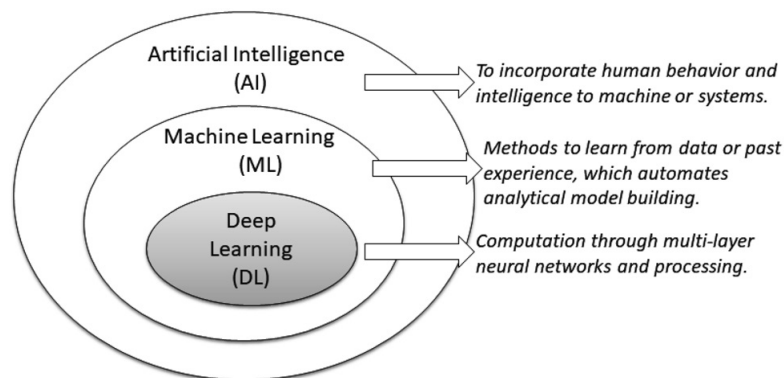


Figura 2.9: Rappresentazione del mondo dell'intelligenza artificiale e delle sue sottobranche: il machine learning e il deep learning. Immagine estratta da [32].

Con il termine machine learning (ML) si intende un metodo di analisi che va ad analizzare anche l'esperienza per favorire il miglioramento della performance dell'algoritmo. Difatti, l'idea alla base del ML è quella di apprendere, da calcoli precedenti, dei dati e degli elementi che possano dare come output del modello delle decisioni affidabili e ripetibili. Il machine learning può essere diviso in tre tipologie [33]:

- *supervised learning*: l'apprendimento supervisionato richiede un set di dati di formazione che comprenda degli esempi in input e i label (etichette) per gli output. Questi valori vengono utilizzati per l'allenamento del modello, il quale

viene poi utilizzato su un set di dati che non sono "etichettati" ottenendo un output veritiero;

- *unsupervised learning*: diversamente dal caso precedente, l'apprendimento non supervisionato non prevede che il modello venga allentato su un set che abbia degli elementi "etichettati". In questo caso il modello viene allenato andando a valutare quali sono gli elementi che devono essere estratti per ottenere un modello che funzioni correttamente anche su altri set di dati;
- *reinforcement learning*: in questo caso non vengono affidate le coppie (input,output) ma vengono valutati l'obiettivo e delle azioni che possono essere compiute, quindi il modello deve raggiungere l'obiettivo utilizzando il principio di tentativi ed errori.

Tra gli strumenti che permettono di utilizzare il machine learning si ritrovano le reti neurali artificiali (Artificial Neural Network – ANN) e il deep learning (DL). Le ANN sono modelli la cui struttura si ispira al processo biologico neurale, attraverso delle rappresentazioni matematiche di unità di elaborazione connesse tra loro che prendono il nome di neuroni. Con lo stesso principio delle sinapsi, ogni connessione tra i neuroni trasmette segnali che sono modulati da pesi, che vengono modificati durante il processo di apprendimento. Generalmente la struttura di una ANN è formata da un layer di input ed uno di output, tra i quali possono esserci un numero variabile di layer nascosti che hanno il compito di apprendere una mappa non lineare che colleghi l'input all'output.

Il DL utilizza delle reti, come le ANN, ma che hanno al loro interno un numero maggiore di layer nascosti che generino un'architettura ramificata. Inoltre, il DL utilizza dei neuroni che non sono semplici come quelli dell'ANN, ma che sono capaci di utilizzare delle operazioni avanzate (ad esempio le convoluzioni) o che siano capaci di essere attivati da numerose funzioni rispetto ad una singola funzione di attivazione come nelle ANN.

Nel caso del deep learning si possono distinguere differenti architetture, tra le quali:

- *convolutional neural network (CNN)*: le CNN sono utilizzate soprattutto nell'ambito delle rappresentazioni a video o nel riconoscimento vocale. Sono caratterizzate dall'essere coinvolte con dataset che non permettono di intercambiare le righe con le colonne, come nel caso delle immagini. L'architettura delle CNN è caratterizzata da una serie di fasi che consentono l'apprendimento gerarchico delle features per definire il task richiesto. La consequenzialità delle features estratte permettono di essere utilizzate come elementi di previsione per input successivi;
- *autoencoders*: questi vengono utilizzati con diverse tipologie di input e sono costituiti da due differenti stadi. Il primo stadio è quello dell'encoding, in

cui vengono estratte le informazioni dall'input, il quale viene compresso in un elemento a bassa dimensione rappresentativa non corrotto dal rumore. La seconda fase, invece, è quella del decoding, la quale consiste nella ricostruzione dell'input originale partendo dalle informazioni carpite nella fase di encoding. Gli autoencoders vengono comunemente utilizzati in algoritmi non supervisionati o per effettuare una riduzione della dimensionalità del dataset in combinazione con altri elementi di apprendimento successivi;

- *generative adversarial neural network (GAN)*: l'idea alla base delle GAN è quella di apprendere una distribuzione di probabilità di un certo dataset e di utilizzarli per generare nuovi campioni di dati. L'architettura di una GAN è caratterizzata dalla presenza di due reti: la prima è una rete di generatori che ha lo scopo di analizzare la distribuzione dell'input generando nuovi esempi; la seconda, invece, è una rete discriminatrice che cerca di distinguere gli elementi reali da quelli artificiali. L'allenamento delle due reti avviene fin tanto che il discriminatore non è più in grado di distinguere gli elementi reali da quelli artificiali.

La rappresentazione di un esempio di struttura di una rete neurale del deep learning è mostrata in figura 2.10.

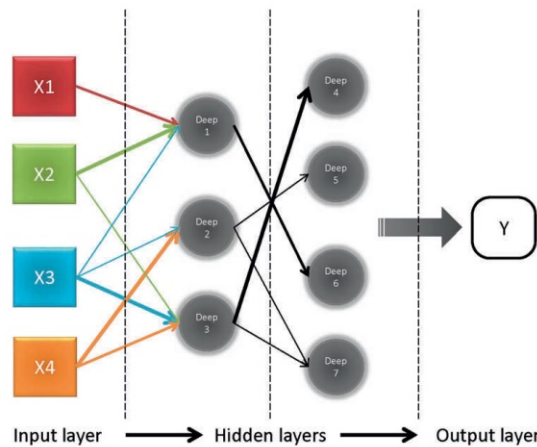


Figura 2.10: Struttura base di una rete neurale nel deep learning. Immagine estratta da [34].

2.3 Generative adversarial network (GAN)

Il concetto di *generative adversarial network* nasce nel 2014 da un'idea di Goodfellow et al. [35] basata su un approccio di apprendimento non supervisionato attraverso l'utilizzo di un modello generativo, il quale ha come scopo quello di trovare una distribuzione del modello ($p_{modello(x)}$) che possa avvicinarsi alla distribuzione definita dagli esempi utilizzati nella fase di allenamento ($p_{data(x)}$).

Dalla schematizzazione dell'architettura della GAN, mostrata in figura 2.11, si nota come quest'ultima sia costituita da due elementi, quali il generatore e il discriminatore, costituiti come reti neurali. Inoltre, dal termine "adversarial", ovvero in contraddizione, si evince come i due elementi siano costituiti per essere in contrasto tra loro. Lo scopo ultimo delle GAN, infatti, è sintetizzabile dall'idea per cui il guadagno di uno è pari alla perdita dell'altro.

Il generatore è un elemento che, prendendo i campioni dallo spazio latente \mathbf{Z} caratterizzata da una distribuzione $p_z(\mathbf{Z})$, genera dei nuovi campioni utilizzando una funzione differenziale $G(z; \Theta_g)$, i quali hanno una distribuzione p_g definiti dai parametri del generatore (Θ_g). Per ottenere questi campioni, il generatore viene allenato utilizzando la backpropagation, ovvero un metodo di stima del gradiente utilizzata dall'algoritmo per ottimizzare il modello e poter calcolare e aggiornare i parametri di rete. Il discriminatore è caratterizzato anch'esso da una funzione differenziale $D(x; \Theta_d)$, dove Θ_d sono i parametri del discriminatore che produce in output la probabilità che gli elementi reali \mathbf{x} provengano dalla distribuzione dei dati dell'addestramento p_{data} . Come detto, lo scopo del generatore è quello di ottenere una distribuzione dei dati che sia paragonabile a quella dei dati di allenamento e questo si raggiunge nel momento in cui il discriminatore non è più capace di distinguere gli elementi generati da quelli reali.

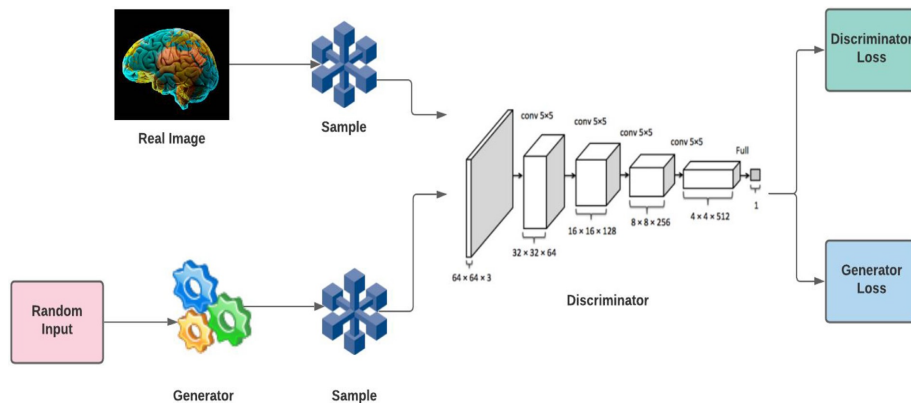


Figura 2.11: Diagramma a blocchi di una generative adversarial network (GAN). Immagine estratta da [36].

Il generatore e il discriminatore sono caratterizzati da una specifica funzione di loss, le quali sono rispettivamente:

$$\begin{aligned} J_G &= -\frac{1}{m} \sum_{i=1}^m \log(D(G(z_i))) \\ J_D &= -\frac{1}{m} \sum_{i=1}^m \log(D(x_i)) - \frac{1}{m} \sum_{i=1}^m \log(1 - D(G(z_i))) \end{aligned} \quad (2.1)$$

La funzione J_G valuta quanto bene il discriminatore riesce ad ingannare il discriminatore e la componente logaritmica rappresenta la probabilità che il discriminatore classifichi come generata un'immagine definita dal generatore. Quindi, il generatore mira a minimizzare questa perdita favorendo dei campioni da esso generato che il discriminatore classifichi come reali (la condizione ideale è che essa tenda ad essere prossima ad uno). Invece, la funzione J_D valuta la capacità del discriminatore di poter classificare correttamente i campioni reali e generati. In questo caso si hanno due componenti logaritmiche: la prima definisce la capacità del discriminatore di classificare correttamente i campioni reali, mentre la seconda definire la sua capacità di classificare i campioni generati come tali. Quindi questa funzione si riduce nel momento in cui si ha una corretta classificazione.

Possiamo, quindi, definire la funzione di ottimizzazione di una GAN come:

$$\min_G \max_D f(D, G) = E_{x \sim p_{data}}[(\log D(x))] + E_{z \sim p_z}[(\log(1 - D(g(z))))] \quad (2.2)$$

dove $D(g_z)$ è la probabilità che il discriminatore identifichi come autentici i dati provenienti dal generatore e $D(x)$, invece, rappresenta la probabilità che il discriminatore identifichi i dati reali come tali.

Le reti GAN vengono utilizzate soprattutto con elementi visivi, quali immagini o video. Nell'ambito dell'analisi delle immagini, tra le applicazioni delle GAN si ricordano:

- *image-to-image translation*: lo scopo di questa applicazione è quello di apprendere la mappatura di un'immagine (X) in modo da poterla trasferire ad una seconda immagine (Y). Il punto fondamentale di questa applicazione è che la modifica dello "stile" dell'immagine non deve minare il contenuto semantico dell'immagine stessa. In questo ambito, possono essere usate differenti reti neurali tra le quali la CycleGAN, dove un esempio del risultato del suo utilizzo è mostrato in figura 2.12 o la Pix2Pix GAN;
- *image super-resolution reconstruction*: ha come scopo quello di ricostruire un'immagine in alta risoluzione partendo da un'immagine a bassa risoluzione, attraverso l'elaborazione di elementi che costituiscono l'immagine. Questi hanno definito diverse modalità d'utilizzo, quali: modello basato sull'interpolazione (il più semplice), modello basato sulla ricostruzione e infine quelli

basati sull'apprendimento. Un'idea è mostrata in figura 2.12 con l'utilizzo della SRGAN;

- *style transformation*: in questo caso si fa riferimento alla possibilità di trasferire stile di un'immagine (X) ad una seconda immagine (Y). Per effettuare questa tipologia bisogna considerare sia il livello semantico sia lo stile del colore dell'immagine. Anche in questo caso, può essere utilizzata una cycleGAN ed un'idea è mostrata in figura 2.12
- *image generation*: l'idea alla base di questa applicazione è quella di andare a realizzare delle immagini realistiche che abbiano al loro interno un contenuto semantico voluto. Alcuni esempi di possibili mezzi per questo scopo sono gli algoritmi SPADE e Pix2Pix. Nella figura 2.12 sono mostrati degli esempi di generazione di immagini di entrambi gli algoritmi;

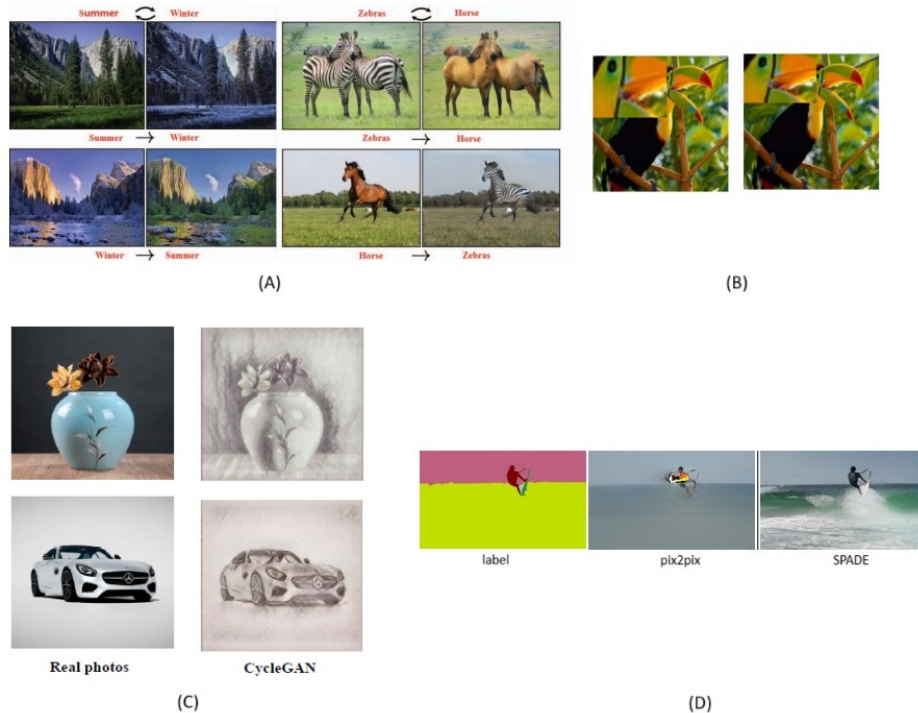


Figura 2.12: Rappresentazioni dei risultati delle diverse applicazioni delle GAN: (A) output di una CycleGAN per la image-to-image translation [37]; (B) output di un algoritmo di image super-resolution reconstruction [38]; (C) output di una CycleGAN per lo style transformation [39]; (D) output di Pix2Pix e dell'algoritmo SPADE per l'immagine generation [40].

2.3.1 Pix2Pix GAN

Il modello Pix2Pix nasce da un'idea di Isola et al. [41], i quali progettaron una conditional generative adversarial network (cGAN) con il proposito di utilizzarla per la image-to-image translation.

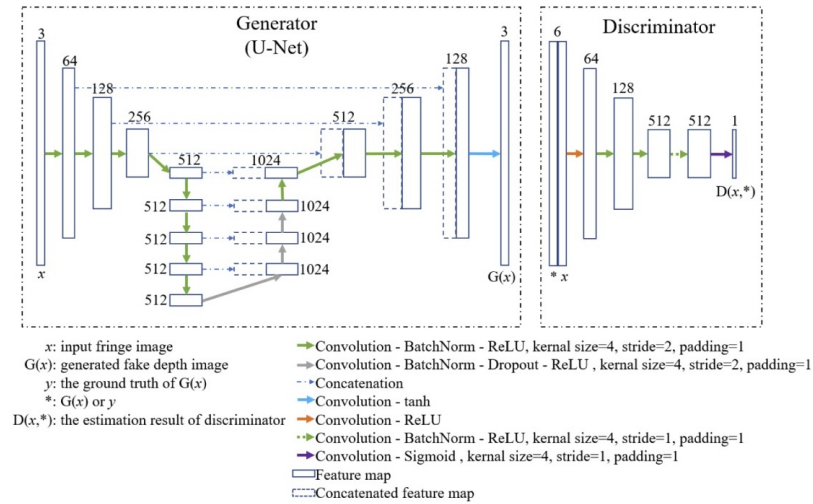


Figura 2.13: Rappresentazione in dettaglio dell'algoritmo Pix2Pix (immagine estratta da [42]).

La struttura di una Pix2Pix GAN, mostrata in figura 2.13, è costituita da un generatore e da un discriminatore, ma diversamente da una normale GAN il generatore è costituito da una U-Net mentre il discriminatore è composto da un discriminatore Markov, caratterizzato dal fatto che, diversamente dai normali discriminatori che hanno in output un solo valore, questo ha un output con valori multipli in quanto la classificazione avviene su ogni singola patch in cui viene divisa l'immagine.

Come detto, il generatore è caratterizzato dall'essere una rete neurale di tipo U-Net, struttura di tipo encoder-decoder, il primo è definito da una serie di fasi di downsampling; mentre il secondo da una fase di upsampling, che ha il compito di riportare l'immagine alle dimensioni originali. Ogni fase è caratterizzata da una serie di convoluzioni, le quali sono mostrate nel dettaglio in figura 2.13. Questo modello è caratterizzato da una funzione di ottimizzazione che richiama quella della GAN ma a cui viene aggiunto un terzo elemento la norma L_1 che migliora la convergenza del modello potendo così migliorare anche le qualità delle immagini generate e accelerando la convergenza.

In questo modo la funzione di ottimizzazione diviene:

$$\begin{aligned} \min_G \max_D f(D, G) = & E_{x \sim p_{data}}[(\log D(x))] \\ & + E_{z \sim p_z}[(\log(1 - D(g(z))))] \\ & + E_{x,y}[(\|x - G(z)\|_1)] \end{aligned} \quad (2.3)$$

2.3.2 SPADE

L'algoritmo SPADE (spatially-adaptive-denormalization) nasce dall'idea di Park et al. [40], con l'intento di utilizzare la sintesi di immagini condizionali, ovvero la conversione di maschere di segmentazione semantica in immagini reali. Questo è un metodo di normalizzazione condizionale, in cui l'attivazione è normalizzata in maniera canalizzata e successivamente modulata, come nella normalizzazione del batch.

L'architettura del blocco SPADE è visualizzabile a sinistra nella figura 2.14.

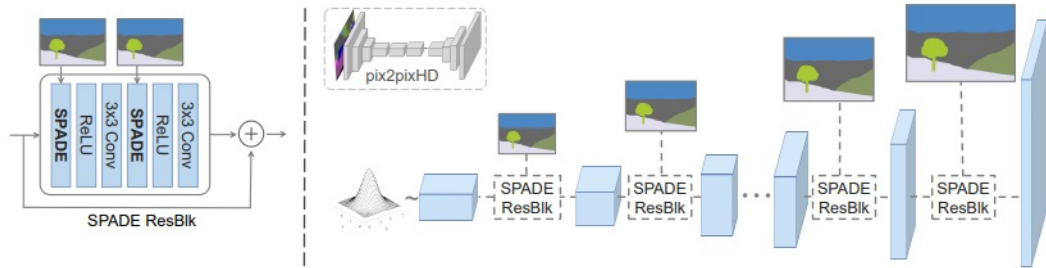


Figura 2.14: Rappresentazione dell'architettura del generatore considerato nel modello SPADE. (Sinistra) Struttura di un blocco residuo; (Destra) Architettura del generatore con al suo interno sono rappresentati i blocchi residui specifici. L'immagine è estratta da [40].

Nella figura di sinistra, possiamo notare come in questo blocco la maschera viene prima proiettata su uno spazio e successivamente convoluta per ottenere i parametri di modulazione β e γ , i quali, diversamente dai precedenti modelli, sono dei tensori con dimensioni spaziali. Questi vengono moltiplicati tra loro e successivamente aggiunti all'attivazione normalizzata elemento per elemento. Invece, nella parte destra della stessa figura 2.14, viene mostrata l'architettura del generatore, che utilizza al suo interno diversi blocchi ResNet con dei layer di sovracampionamento, i quali apprendono i parametri di modulazione attraverso il modello SPADE. Al termine di ogni blocco residuo, la maschera semantica viene sottocampionata per poter corrispondere alla risoluzione spaziale.

Tra gli aspetti positivi di questo modello possiamo definire la possibilità di preservare al meglio il contenuto semantico dell'immagine rispetto, ad esempio all'uso della

pix2pix, questo perché la maschera di segmentazione viene modulata spazialmente senza che ci sia normalizzazione.

2.4 Applicazioni e approcci del ML in patologia

Nell'ambito della patologia, negli ultimi anni sono stati apportati dei metodi che permettono di utilizzare il machine learning per diversi scopi, quali la segmentazione delle cellule, la diagnosi o l'utilizzo del ML per la predizione di possibili gradi di patologia all'interno dell'immagine. Generalmente il workflow che viene seguito è il seguente (mostrato in figura 2.15), il quale consta di cinque passaggi:

1. ottenimento delle immagini: in questa fase si fa riferimento all'ottenimento delle WSI;
2. preprocessing o scrub, con la stain normalization e la divisione in patch;
3. l'esplorazione, ovvero la scelta del modello più consono al compito da svolgere;
4. l'applicazione del modello scelto;
5. l'interpretazione del risultato ottenuto.

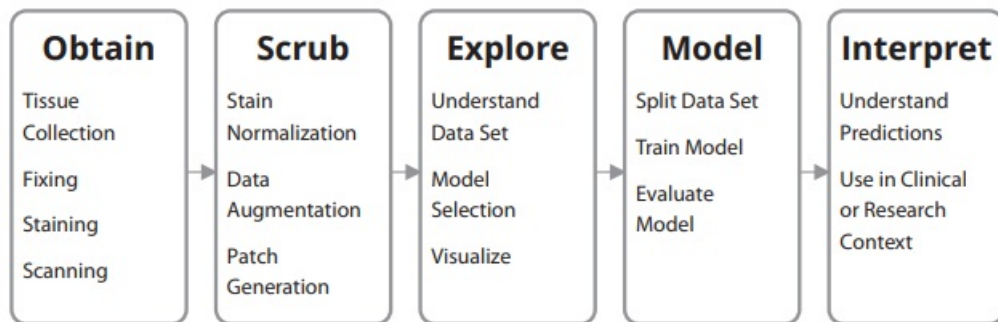


Figura 2.15: Workflow utilizzato in caso di deep learning in patologia. Immagine estratta da [43].

Nel momento in cui si parla di utilizzare un metodo di machine learning nell'ambito patologico, si fa riferimento all'utilizzo delle WSI, le quali hanno alcune problematiche di utilizzo. Il primo problema è relativo alle loro dimensioni (ordine dei gigapixel), le quali determinano la necessità di dividerle in patch per poter essere valutate correttamente. Una seconda problematica è quella relativa alla necessità di avere delle "label", che permettano di utilizzare dei metodi di ML di tipo supervisionata. In questo caso entra in gioco la figura del patologo, che deve

definire le caratteristiche prime degli elementi della WSI. Inoltre, l'ottenimento delle stesse immagini è soggetto ad artefatti che possono modificare il risultato del modello utilizzato. Per questo motivo nella seconda fase si predilige una fase di stain normalization o di data augmentation. La prima è una fase che permette la normalizzazione del colore, il quale è causa di modifica del risultato del modello. Questo avviene in quando i coloranti utilizzati nei vari istituti non sono uguali tra loro e di conseguenza si possono creare delle differenze nel momento in cui si creano dei dataset con immagini provenienti da differenti istituti. Allo stesso modo si hanno problematiche relative alle quantità esigue di elementi che possono comporre un dataset, dovendo quindi utilizzare una fase di data augmentation che permetta di andare ad aumentare le immagini che costituiscono il dataset e che possano evitare anche una condizione di overfitting del modello.

Generalmente con l'utilizzo delle WSI si fa riferimento alle GAN (generative adversarial neural network) e alle sue varianti. Difatti, tra le soluzioni proposte da Morrison et al. [43], nell'attuale stato dell'arte, si fa riferimento a tre possibili approcci:

- il primo è quello di considerare il trasferimento dello stile di stain (style–stain transfer), ovvero alla possibilità di poter trasferire uno stile preso da una sorgente ad un'altra immagine, che sia il target. Con il termine stile si fa riferimento non solo al colore ma anche alla disposizione degli elementi che sono presenti all'interno dell'immagine (contenuto semantico);
- utilizzo della Pix2Pix per la image–to–image translation (o traduzione da immagine ad immagine). Questa è un'estensione delle conditional GANS, che permette di allenare un modello senza specificare una determinata funzione di loss tra il dominio dell'immagine sorgente e quella target. L'unico elemento che serve a questa è la coppia di immagini (sorgente,target);
- utilizzo delle CycleGAN, dove diversamente dal caso delle Pix2Pix non necessitano di una coppia di immagini ma hanno bisogno solo di immagini sorgenti.

2.4.1 Open challenge

Nel momento in cui si vogliono ottenere dei risultati veritieri, nel momento in cui si voglia utilizzare il ML nell'ambito patologico, bisogna cercare di superare le varie limitazioni che sono attualmente considerate come open challenge.

Quelle poste in esame da Cui et al. nel suo studio [34] sono:

- la necessità di ottenere una standardizzazione ed una normalizzazione delle WSI. La necessità di ottenere un protocollo standardizzato è alla base della possibilità di ridurre o eliminare completamente gli errori di sistema e gli errori

casuali derivanti dalla diversa strumentazione. La nascita di uno standard permetterebbe di ottenere un dataset con immagini provenienti da diversi istituti. Un primo standard è stato apportato dalla DICOM (digital imaging and communications in medicine), il quale ha affrontato anche il problema relativo ad immagini radiologiche;

- la necessità di un team interdisciplinare, dove la figura del patologo ha un ruolo critico, in quanto allo stesso non è richiesta la sola conoscenza medica ma anche una conoscenza in ambito statistico e del maneggiamento dei dati, in modo da poter ridurre il divario presente tra la medicina e l'intelligenza artificiale;
- la necessità di superare le limitazioni degli attuali hardware. Attualmente l'utilizzo delle WSI è limitato a causa delle loro dimensioni e dal bisogno di un'ingente quantità di memoria disponibile. Inoltre, lo stesso processo grafico alla base dell'analisi delle WSI risulta essere attualmente limitato. Quindi, nell'analisi delle WSi si ha bisogno di una strumentazione idonea;
- il problema etico relativo al trasferimento dei dati sensibili.

Altre proposte, relative alle open challenge, sono state presentate da Janowczyk et al. [44] e concernano per lo più l'ambito delle WSI e dell'allenamento e sono:

- scelta del giusto ingrandimento delle WSI;
- gestione degli errori relativi ai "label": questo problema si presenta nel momento in cui viene utilizzato un modello supervisionato, poiché questo richiede che alle immagini siano associate delle etichette che solitamente sono definite dai patologi. Queste, però, non sono effettuate manualmente dai patologi e di conseguenza sono soggette ad una mancata precisione, magari nel momento in cui sono effettuate su immagini con un basso ingrandimento. Inoltre, un errore a livello delle label può causare un errore nell'allenamento del modello e di conseguenza a dei falsi positivi;
- identificazione di un set di addestramento corretto: in questo caso si fa riferimento al corretto numero di patch in cui dividere il dataset e le dimensioni corrette per l'allenamento del modello. Inoltre, risulta essere indispensabile capire bene quale modello utilizzare e di conseguenza quali elementi (come ad esempio la funzione di loss o il numero di epoche) considerare.

Capitolo 3

Metodi

3.1 Presentazione del Dataset

Il dataset utilizzato per questo lavoro di tesi proviene da un lavoro precedente. Inizialmente il dataset consisteva in 331 WSI (whole slide images) provenienti dal Vancouver Prostate Center, le quali sono ottenute attraverso uno scanner SCN400 con una risoluzione di 40x ed una colorazione con H&E (ematossilina–eosina). Tali immagini prevedevano al loro interno delle annotazioni ottenute grazie l’ausilio di due patologi, le quali definivano la classificazione dei tessuti rispettivamente in cinque classi: benigno, Gleason 3, Gleason 4, Gleason 5 e non marcato.

Le fasi di pre–processing utilizzate nel precedente lavoro di tesi sono mostrate in figura 3.1 e comprendono:

- creazione del ground truth label: questo passo prevedeva l’utilizzo dell’algoritmo STAPLE (Simultaneous Truth And Performance Level Estimation), il quale ha permesso di correggere alcuni errori relativamente ad alcune annotazioni, che presentavano dei contorni aperti e non pienamente marcati. Quindi lo scopo di questo passaggio era il completamento delle annotazioni e la corretta assegnazione della classificazione;
- ottimizzazione e normalizzazione delle WSI con conseguente generazione delle maschere dei bianchi (zone appartenenti al vetrino) e dei nuclei con i relativi bordi;
- integrazione delle mappe semantiche: le mappe semantiche sono pari all’integrazione delle maschere dei bianchi e dei nuclei all’interno dei ground truth;
- creazione delle mappe di istanza: queste sono delle mappe che contengono al loro interno i bordi dei nuclei e dei bianchi.

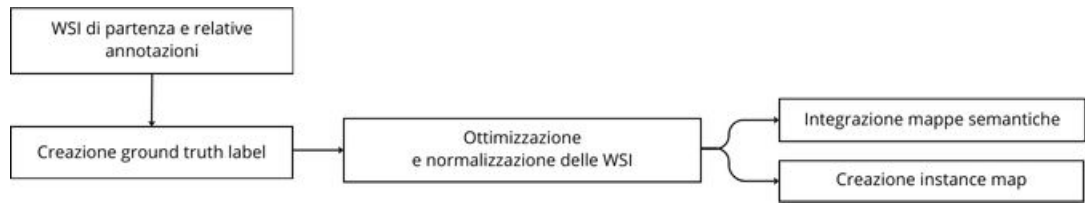


Figura 3.1: Processo di pre-processing delle WSI nel precedente lavoro di tesi.

In figura 3.2 è mostrata la stessa WSI nel passaggio che porta alla sua ottimizzazione e normalizzazione, la quale viene descritta come forma di pre-processing.

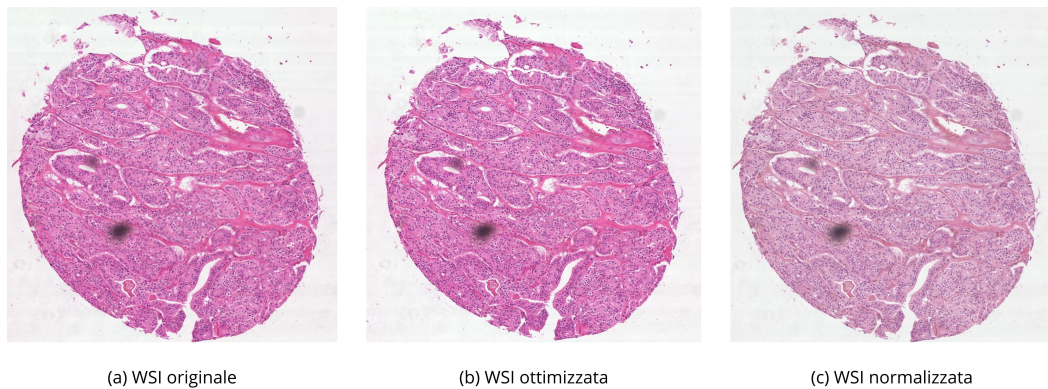


Figura 3.2: Rappresentazione del processo di ottimizzazione e normalizzazione delle WSI. Si definiscono rispettivamente: (A) WSI originale, (B) WSI ottimizzata e (C) WSI normalizzata.

3.2 Pipeline

La pipeline del lavoro svolto è mostrata in figura 3.3. Possiamo racchiudere il lavoro in quattro macrofasi:

- fase di patch extraction, in cui le WSI vengono divise per poterle utilizzare come input per l'algoritmo;
- creazione del dataset, in cui si analizza la numerosità delle patch e si dividono in training e test set;
- implementazione dei dati per l'algoritmo SPADE;
- valutazione delle metriche no-reference e della texture analysis.

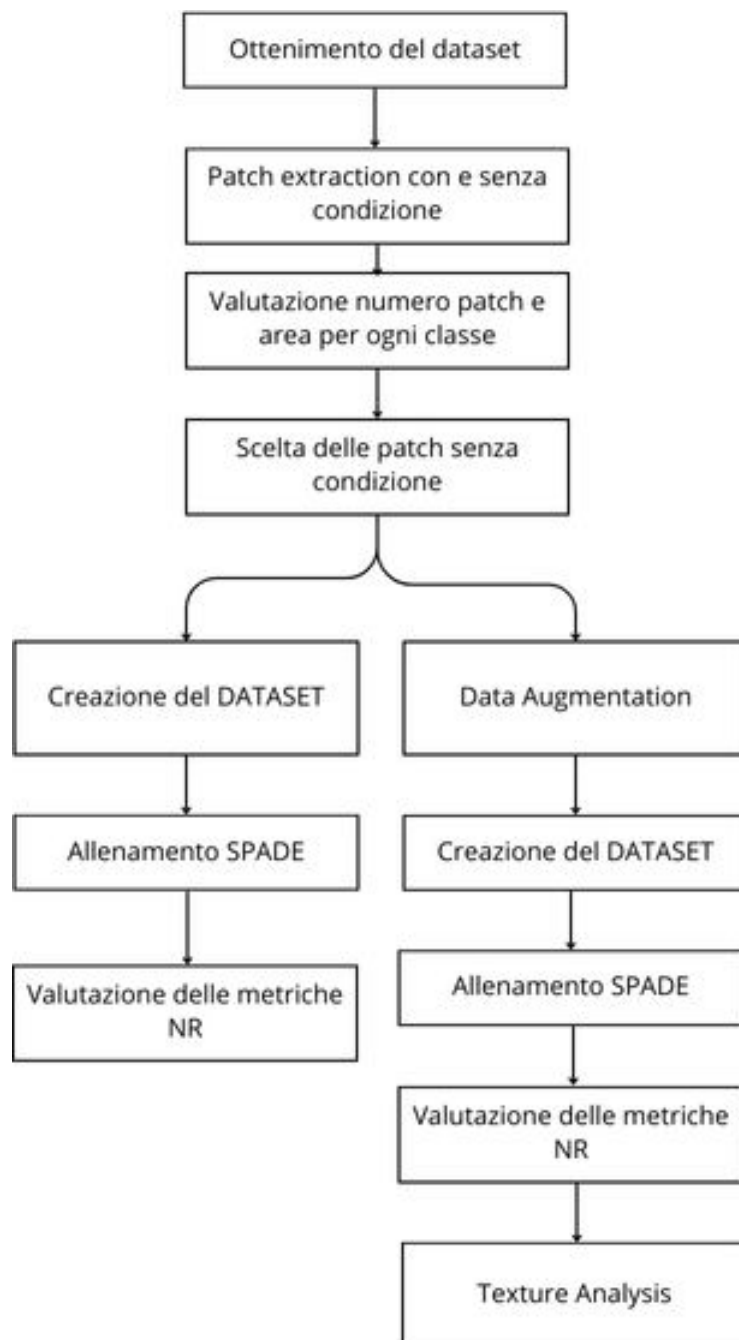


Figura 3.3: Pipeline del lavoro svolto.

3.3 Patch extraction: metodi e valutazione

Come già accennato nel paragrafo 2.1.1, la fase di patch extraction è alla base dell'utilizzo delle WSI come input di un modello generativo. Tale necessità nasce dal problema relativo alle dimensioni delle WSI le quali, come detto precedentemente, sono nell'ordine dei gigapixel. Inoltre, la possibilità di dividere le WSI in patch permette di andare a valutare il livello informativo, potendo così eliminare quelle che sono delle informazioni ridondanti.

Le caratteristiche che devono essere valutate in fase di patch extraction sono le dimensioni e l'ingrandimento scelto, i quali sono fondamentali per valutare il livello informativo contenuto nell'immagine stessa. In particolare, nel caso corrente sono state scelte delle patch quadrate con dimensioni 512x512 ed un ingrandimento 20x. Inoltre, come definito nel paragrafo 2.1.1, ci sono differenti possibilità di estrazione delle patch in relazione alla tipologia di divisione che si vuole effettuare. Nel momento in cui tali patch vengono utilizzate per un modello generativo, la soluzione ottimale è quella che prevede lo scorrimento della finestra con la presenza di un overlap, ovvero una sovrapposizione. Nel caso corrente è stato utilizzato un overlap pari ad un quarto delle dimensioni delle patch (128).

L'estrazione delle patch da utilizzare è avvenuta andando ad effettuare questo processo in due differenti modalità: la prima, come nel precedente lavoro di tesi, presentava una condizione di estrazione, ovvero che nella patch doveva essere presente almeno il 90% di una determinata label; la seconda, invece, prevedeva l'estrazione senza l'imposizione di alcuna condizione. Per la valutazione della modalità di estrazione da utilizzare è stata analizzata la numerosità delle sole label che permettessero di valutare la presenza o meno della condizione cancerogena, ovvero: benigno, Gleason 3, Gleason 4 e Gleason 5.

I valori ottenuti per entrambi le condizioni di estrazione, divisi per label, sono mostrati in tabella 3.1.

Label	Condizione al 90%	Senza condizione
Benigno	2557	7826
Gleason 3	15442	33751
Gleason 4	527	1466
Gleason 5	3668	14431
Totale	22194	57474

Tabella 3.1: Tabella riassuntiva della numerosità delle patch per le singole label analizzata per entrambi i metodi di estrazione.

Per valutare correttamente quale condizione fosse preferibile allo scopo, sono state valutate anche le percentuali delle aree medie delle varie label rispetto all'area della patch totale (ovvero 262144 pixel). Le aree sono state considerate solo quando presente all'interno della patch. I dati sono mostrati in tabella 3.2.

Label	Condizione al 90%	Senza condizione
Benigno	79.55%	41%
Gleason 3	95.32%	57.39%
Gleason 4	81.90%	45.82%
Gleason 5	35.36%	21.78%

Tabella 3.2: Tabella riassuntiva delle aree (in valore percentuale) delle patch per le singole label analizzata per entrambi i metodi di estrazione.

Per rendere l'algoritmo più stabile è stato deciso di utilizzare l'estrazione priva di condizionamento, valutando come la numerosità delle patch e le aree d'analisi fossero bilanciate tra loro. In questo modo si è poi proceduto al primo allenamento.

3.4 Creazione del dataset

Il dataset viene creato partendo dalle patch senza condizione, andando a suddividere queste ultime in Training e Test set. In particolare la suddivisione avviene in maniera randomica, considerando che il 70% delle patch vada a costituire il Training Set, mentre il restante 30% il Test Set. In seguito alla suddivisione si è ottenuto un dataset contenente 65329 elementi per il Training Set e 27998 elementi per il Test Set. Per ogni set, le patch erano presenti come immagini originali, mappe di segmentazione e mappe d'istanza. Un esempio di elementi prelevati dal Training set è mostrato in figura 3.4.

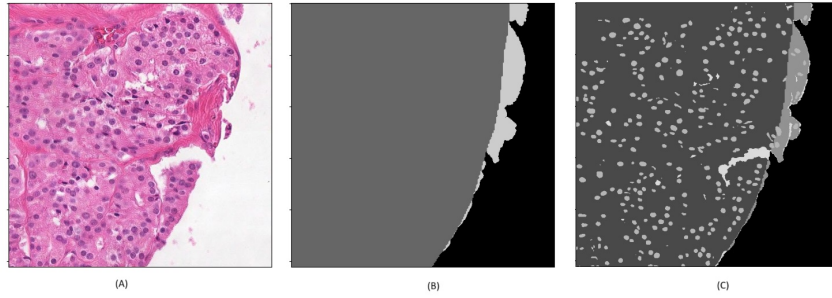


Figura 3.4: Esempio degli elementi presenti nel Training Set, dove si ha: (A) patch, (B) mappa di segmentazione e (C) mappa d’istanza.

Invece, la distribuzione delle classi all’interno dei due set è mostrata in tabella 3.3.

Classe	Training Set	Test Set
Sfondo	11504	4895
Benigna	5432	10119
Gleason 3	23632	465
Gleason 4	1001	4337
Gleason 5	10094	5788
Altro	13666	5788
Totale	65329	27998

Tabella 3.3: Distribuzione delle classi all’interno del Training e del Test set.

3.5 Allenamento del modello

Per la generazione delle immagini è stato utilizzato l’algoritmo SPADE, analizzato in dettaglio nel paragrafo 2.3.2. Sono stati effettuati due allenamenti. Il primo, con il dataset descritto nel paragrafo precedente ed un secondo algoritmo con un secondo dataset, ottenuto dopo una fase di data augmentation solo per le classi del Gleason 4 e della classe benigna, le quali risultano essere sottostimate come output del modello.

I parametri impostati per l’allenamento del modello, coincidono in entrambi gli allenamenti, tranne per il numero di epoche considerate. Difatti vengono analizzate trenta epoche totali per il primo allenamento e cinquanta epoche totali per il secondo. In particolare, durante il secondo allenamento, le valutazioni (mostrate

nel capitolo 4) sono state effettuate al raggiungimento delle trenta, delle quaranta e delle cinquanta epoche. Gli altri parametri considerati sono:

- batch size: iperparametro che definisce il numero di campioni su cui lavorare prima di aggiornare i parametri del modello interno. Nel modello questo valore è stato impostato ad 1;
- niter: numero di epoche considerate prima di aggiornare il learning rate. Questo parametro è stato impostato a 25 per il primo allenamento e a 45 per il secondo;
- niter decay: numero di epoche per far decadere linearmente il learning rate a zero. Questo è stato impostato a 5. In particolare la somma tra il niter e il niter decay, va a definire il numero totale di epoche di allenamento;
- learning rate: parametro di regolazione che determina la dimensione del passo ad ogni iterazione mentre si muove verso il minimo della funzione di loss. Questo è stato impostato a 0,002;
- β_1 e β_2 : sono i parametri dell'algoritmo di ottimizzazione, ovvero l'algoritmo ADAM (ADAPtive Moment estimation) e sono stati scelti rispettivamente come 0,9 e 0,9999.

3.6 Data augmentation

Come accennato nel paragrafo precedente, in seguito ai risultati ottenuti dal primo allenamento, è stata considerata necessaria una fase di data augmentation a causa di una sotto rappresentazione della classe Gleason 4 e della classe Benigna. Il processo di data augmentation è un metodo utilizzato per l'aumento degli elementi presenti all'interno di un dataset, attraverso la creazione di nuovi campioni partendo dalle immagini presenti all'interno del dataset.

Tale fase di data augmentation è stata effettuata due volte per ottenere una numerosità, relativa a queste classi, che permettessero di non essere sottostimate. Prima di procedere al data augmentation, le immagini sono state divise in relazione alla presenza del label in analisi, in modo che la numerosità delle altre classi rimanga costante, mentre quella delle classi sotto rappresentate aumenti. Alcune tra le modifiche che sono state effettuate si ricordano: il flip orizzontale, il flip verticale, modifica del contrasto e della luminosità, rotazione e la trasposizione.

Successivamente al primo data augmentation, la classe benigna risulta essere bilanciata rispetto alle altre, mentre la classe Gleason 4 è risultata bilanciata solo dopo la seconda volta che è stato effettuato il processo di aumento dei dati.

I valori ottenuti nelle due fasi di data augmentation sono mostrati in tabella 3.4.

Numerosità	Prima volta	Seconda volta
Sfondo	11508	11508
Benigna	10819	10819
Gleason 3	23651	23692
Gleason 4	1997	7960
Gleason 5	10121	10123
Altro	13666	13666
Totale	71762	77768

Tabella 3.4: Numerosità delle patch rispetto alle varie classi successivamente al primo e al secondo data augmentation.

In figura 3.5 possiamo vedere degli esempi di modifiche effettuate sulla stessa patch.

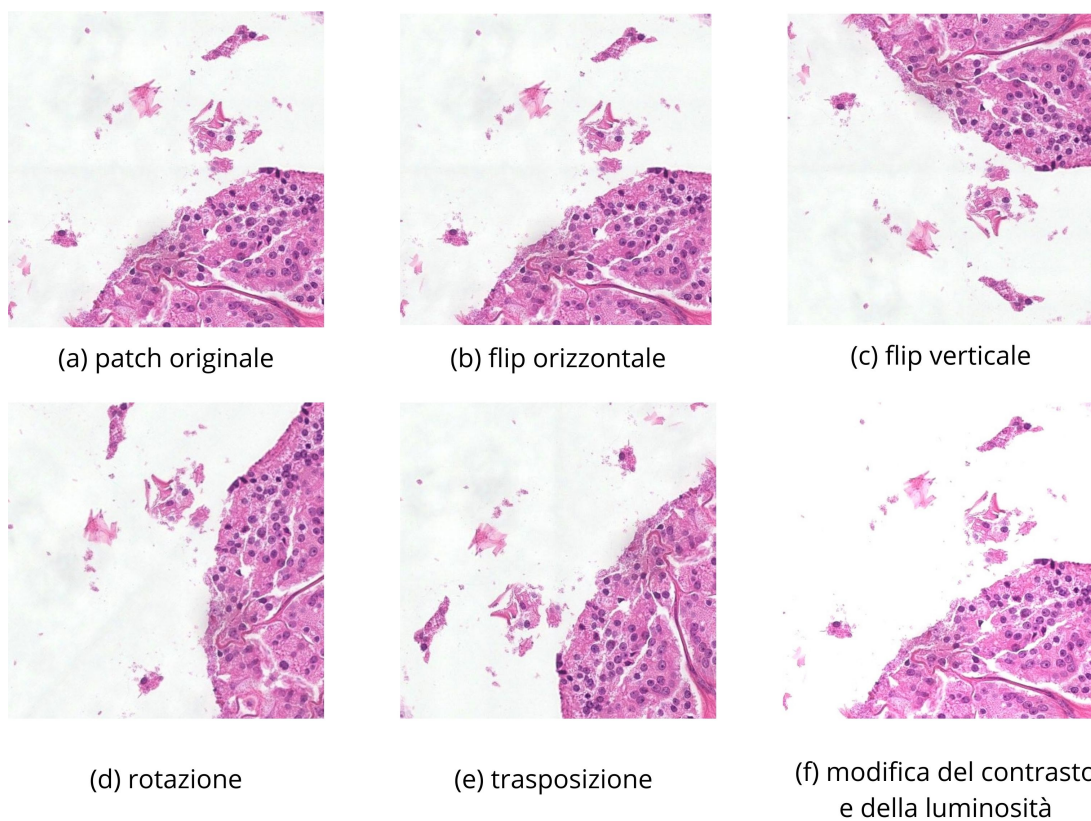


Figura 3.5: Rappresentazione delle diverse modifiche effettuate nel processo di data augmentation.

3.7 Valutazione delle metriche no reference (NR)

Nel momento in cui si vuole valutare la qualità di un'immagine generata, bisogna far riferimento ad alcune features o parametri relativi alla qualità ed è in questo contesto che vengono definite le metriche [45]. In particolare le metriche possono essere divise in tre differenti classi:

- *full-reference (FR)*: in questo caso l'immagine originale viene considerata come riferimento, quindi tali metriche si basano su un confronto con tale riferimento;
- *reduced-reference (RR)*: in questo caso non bisogna avere accesso alle immagini originali, ma basta avere accesso a delle features o delle caratteristiche che vadano a definire le immagini originali;
- *no-reference (NR)*: in questo caso non si utilizza nessun tipo di riferimento delle immagini originali, ma si considerano le sole informazioni provenienti dall'immagine di cui si dispone.

In questo studio si farà riferimento alle ultime metriche definite, ovvero quelle no-reference. L'utilizzo delle metriche NR per la valutazione della qualità dell'immagine (IQA – Image quality assessment) ha, in alcuni casi, come spazio di lavoro quello della valutazione delle distorsioni presenti nelle NSS, ovvero le natural scene statistics, le quali possono essere definite come le caratteristiche statistiche che permettono di definire la naturalezza di un'immagine, ovvero la sua mancata provenienza da elementi che l'abbiano generata artificialmente. In questo caso l'utilizzo di metriche che fanno riferimento a questo aspetto viene considerato in quanto lo scopo principale di questo studio è la possibilità di valutare un algoritmo che permetta di creare elementi che siano tanto più simili a quelli reali.

La valutazione del livello di "naturalezza" è stata effettuata attraverso l'analisi di due metriche, quali: la Blind Referenceless Image Spatial Quality Evaluator (BRISQUE) e la Natural Image Quality Evaluator (NIQE).

Il Blind Referenceless Image Spatial Quality Evaluator (BRISQUE) [46] è un modello di metriche no reference, che utilizza uno spazio NSS con coefficienti di luminescenza localmente normalizzati e attraverso i parametri del modello permette di valutare la "naturalezza" dell'immagine. Il valore di tale metrica può raggiungere un valore compreso tra 0 e 100 a cui ad un valore minore si associa un andamento più naturale dell'immagine.

Invece, il Natural Image Quality Evaluator (NIQE) [47] è un modello che permette di estrarre delle features in forma vettoriale da un'immagine e di adattarli ad un modello gaussiano multivariato (MVG). La qualità dell'immagine viene valutata andando a confrontare la distanza tra il modello ottenuto dall'immagine in esame e quello ottenuto da un corpo di immagini non contaminate. Questo modello non

tiene conto delle percezioni soggettive diversamente dal modello BRISQUE, ma come quest'ultimo si associa una migliore qualità dell'immagine a bassi valori del risultato.

3.8 Valutazione texture analysis

La texture analysis è una branca dell'imaging processing digitale che permette la valutazione della caratterizzazione di un'immagine, Viene utilizzata soprattutto per poter delineare il contenuto dell'immagine attraverso degli elementi che valutino le sue caratteristiche. E' un elemento di valutazione utilizzato nel momento in cui si è in presenza di casi di segmentazione, classificazione e generazione semantica, come nel caso in esame.

Macroscopicamente l'analisi della tessitura può essere effettuata andando ad analizzare due gruppi differenti di features:

- descrittori del primo ordine, i quali sono basati sull'utilizzo dell'istogramma della scala dei grigi;
- descrittori del secondo ordini, i quali sono basati sull'utilizzo della matrice delle co-occorrenze della scala dei grigi (GLCM – gray level co-occurrence matrix).

In questo lavoro vengono analizzati contemporaneamente entrambe le classi di descrittori. In particolare, l'analisi viene effettuata per ogni singola label andando a valutare la differenza tra le immagini reali e quelle generate. La valutazione delle diverse label, viene effettuato andando a valutare i descrittori solo nella porzione dell'immagine che abbia come etichetta quella analizzata. Questo passaggio risulta essere necessario per poter poi effettuare un confronto tra le categorie di immagini (reali e sintetizzate).

3.8.1 Descrittori del primo ordine

Come definito nel paragrafo precedente, i descrittori del primo ordine sono quelli che valutano le features partendo dall'istogramma della scala dei grigi. Questo viene valutato andando ad effettuare il conteggio dei pixel che abbiano un determinato valore di grigio nella scale. In tal senso, si ricorda che il valore minimo è lo zero e rappresenta il colore nero mentre il valore massimo è definito in relazione al numero di bit di cui è composta l'immagine, valutato come 2^{bit} , e rappresenta il colore bianco. Una degli svantaggi principali di questa classe di descrittori è la perdita dell'informazione spaziale, in quanto quella che forniscono è una sintesi delle informazioni relative alle immagini.

I descrittori analizzati, con le relative definizioni, sono definiti nell'appendice A.

3.8.2 Descrittori del secondo ordine

I descrittori del secondo ordine, invece, permettono di analizzare le caratteristiche dell'immagine andando a valutare quella che prende il nome di matrice delle co-occorrenze della scala dei grigi (GLCM). Questa, diversamente dall'istogramma, non dà solo una sintetica informazione delle caratteristiche ma permette di mantenere l'idea di spazialità dei pixel. La GLCM si valuta andando a costruire un'immagine quadrata che abbia i valori dei pixel sia sulle righe (il cui numero è pari al numero massimo che un pixel può raggiungere sulla scala dei grigi) sia sulle colonne; sull'intersezione, invece, si calcola quante volte un pixel con un determinato livello di grigio si trova vicino ad un altro pixel con un determinato valore. Tale coppia di valore può essere valutata in qualunque direzione, andando a considerare un determinato angolo di inclinazione (ad esempio se l'angolo considerato è pari a zero, allora si va a valutare tale elemento orizzontalmente). Inoltre, tale direzionalità non subisce l'effetto del verso in quanto la GLCM è una matrice simmetrica.

Come detto un elemento a favore di queste features è la spazialità che viene mantenuta ma al contempo ha tra gli svantaggi:

- la possibilità di avere diversi valori rispetto all'angolo considerato;
- si valutano solo gli elementi vicini tra loro (occorrenza), rendendosi complicato il suo utilizzo nel momento in cui si considerano degli elementi che abbiano delle sfumature di grigio molto simili.

I descrittori analizzati, con le relative definizioni, sono definiti nell'appendice B.

Capitolo 4

Risultati

Nel seguente capitolo vengono mostrati i risultati ottenuti nel seguente lavoro di tesi. Ad un primo approccio visivo, possiamo valutare una certa coerenza tra le immagini generate e quelle reali, come mostrato in figura 4.1.

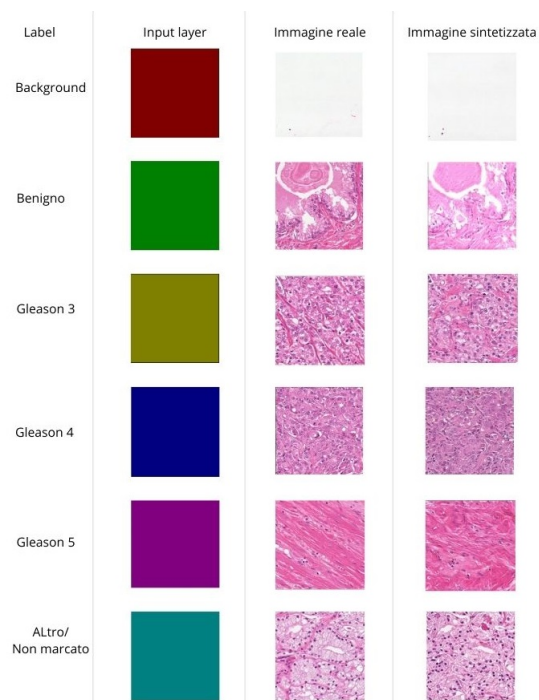


Figura 4.1: Rappresentazione delle immagini reali e sintetizzate, con i relativi input layer, per le sei label: Background, Benigno, Gleason 3, Gleason 4, Gleason 5 e Non marcato. Le immagini sono state ottenute in seguito al secondo allenamento con il numero massimo di epoche (50).

4.1 Metriche No–Reference

In questo paragrafo verranno trattati i risultati dell’analisi metrica ottenuti per i due allenamenti.

Si analizzano inizialmente i risultati ottenuti dal primo allenamento, svolto con le classi "benigna" e "gleason 4" in forma sottostimata rispetto alle altre label. Gli elementi di valutazione considerati sono le sole metriche metriche no–reference, ovvero BRISQUE e NIQE, effettuando il confronto tra le immagini sintetizzate e quelle reali. Tale valutazione è stata effettuata in due differenti modalità: la prima senza considerare la distinzione tra le varie label e la seconda andando a considerare la distinzione tra le label. I dati ottenuti sono mostrati in tabella 4.1 e mostrati graficamente, attraverso l’uso di istogrammi, in figura 4.2.

Immagini	BRISQUE	NIQE
Sintetiche	22.70 ± 17.31	4.55 ± 1.53
Reali	11.68 ± 10.08	3.40 ± 1.02

Tabella 4.1: Confronto delle metriche no–reference per il primo allenamento: valutazione senza distinzione tra le label. I risultati sono mostrati come media \pm deviazione standard.

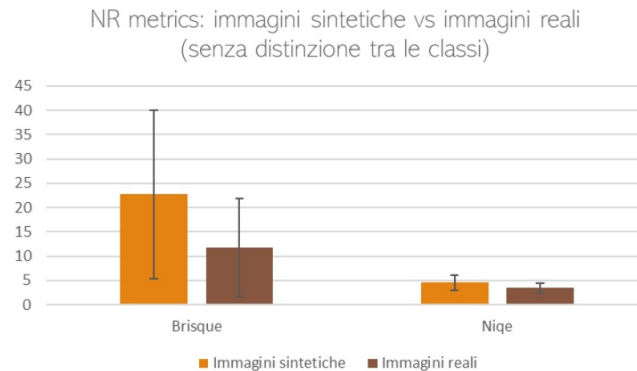


Figura 4.2: Il grafico mette in relazione i dati delle metriche no–reference ottenuti per il primo allenamento, senza distinzione tra i vari label.

Possiamo notare come nel caso della metrica NIQE i valori delle immagini generate non si discostano molto rispetto a quelle reali; lo stesso non si può dire della metrica BRISQUE, la quale risulta essere più inclusiva dell’altra. Possiamo, però, notare che il valore della metrica BRISQUE sia comunque basso rispetto al

massimo ottenibile (100) e che la qualità delle immagini può essere considerata sufficiente.

Andando, invece, a valutare le stesse metriche per le varie label, sono stati ottenuti i risultati mostrati in tabella 4.2.

Metrica	Immagini sintetiche			
–	Ben	G13	G14	G15
BRISQUE	27.96±19.59	19.23±12.67	14.60±12.12	25.40±20.07
NIQE	5.21±2.36	4.46±0.85	3.98±0.90	4.49±1.79
Metrica	Immagini reali			
–	Ben	G13	G14	G15
BRISQUE	9.61±7.38	9.99±8.06	11.20±6.90	14.01±11.98
NIQE	2.98±0.74	3.36±0.83	3.62±0.82	3.53±1.23

Tabella 4.2: Confronto delle metriche no-reference per il primo allenamento: valutazione con distinzione tra le label. I risultati sono mostrati come media \pm deviazione standard.

Valutando la metrica BRISQUE, si nota come l'unica label ad avere un buon punteggio sia la Gleason 4, in cui la differenza dei valori medi tra il valore per le immagini reali e quelle sintetizzate è di 3.40 differentemente dalle altre label che assumono una differenza nettamente maggiore, valutata mediamente e pari a 12.99. Possiamo, quindi, valutare come questo allenamento ha presentato una semplicità di valutazione di tale label rispetto alle altre. Diversamente nell'analisi della metrica NIQE, la differenza media risulta essere proporzionalmente minore rispetto alla metrica BRISQUE, ma si può notare come anche in questo caso la label rappresentata più realisticamente sia la Gleason 4. Allo stesso tempo, questa semplicità potrebbe essere correlata a due elementi che caratterizzano questa label, ovvero: la minore numerosità rispetto alle altre e la minore area occupata sulla patch rispetto alle altre. Come detto nel paragrafo 3.6, questa label insieme alla label "benigna" sono state aumentate a causa di questa discrepanza riscontrata, difatti abbiamo che anche la label benigna è sottostimata rispetto alle altra ma che ha, allo stesso tempo, un'impronta maggiore della Gleason 4, come mostrato in tabella 4.3. Quest'analisi ha fatto sì che venisse riscontrata la necessità di ampliare la numerosità delle due categorie.

Label	Numerosità (%)	Area (%)
Sfondo	57.91	20.53
Benigna	19	7.8
Gleason 3	55.86	32
Gleason 4	2.96	1.34
Gleason 5	77.75	16.89
Non marcato	61.99	21.44

Tabella 4.3: Valutazione percentuale dei valori medi della numerosità delle label sul numero totale di immagini e sull'area delle singole label rispetto all'area totale della patch. I valori sono ottenuti considerando il primo allenamento effettuato.

La rappresentazione grafica dei dati in tabella 4.2 è mostrata in figura 4.3, ottenuta tramite istogrammi.

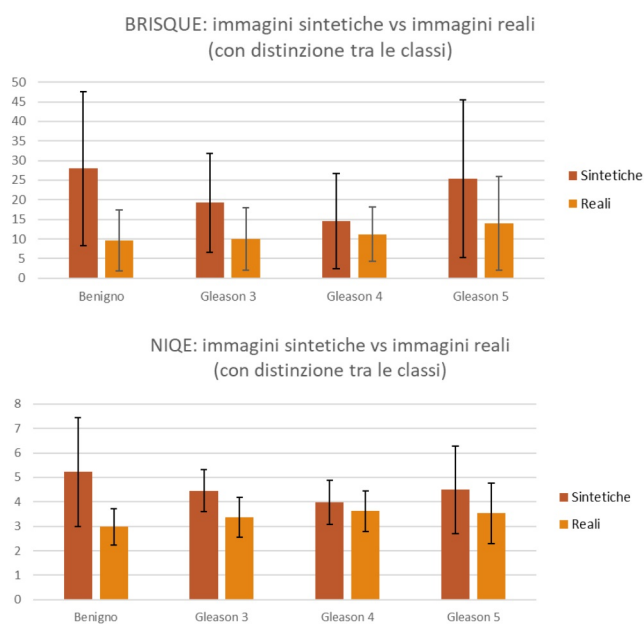


Figura 4.3: Il grafico mette in relazione i dati delle metriche no-reference ottenuti per il primo allenamento, con distinzione tra i vari label.

In seguito al secondo allenamento le immagini sintetizzate, come detto nel paragrafo 3.5, sono state valutate in differenti momenti dell'allenamento stesso, ovvero: al raggiungimento delle trenta epoche, delle quaranta epoche e infine delle cinquanta epoche. Per questo motivo l'analisi dei risultati verrà effettuata andando a considerare i tre differenti momenti. Questa scelta è stata effettuata per poter valutare quale fosse il numero idoneo di epoche per evitare che il modello si portasse ad una condizione di overfitting. Inoltre, come nel primo allenamento, le metriche sono state valutate sia per l'immagine nel complesso sia per le singole label. I dati che sono stati riportati sono mostrati rispettivamente nelle tabelle 4.4 e 4.5. Come mostrato in tabella 4.4, possiamo notare come il valore della metrica BRISQUE tenda ad aumentare con il numero delle epoche ma che al contempo il valore della deviazione standard sia nettamente migliore nel caso di immagini sintetizzate dopo 50 epoche, potendo così valutare con il fitting del modello sia migliore in questo caso. Lo stesso andamento può essere valutato per la metrica NIQE, ma con una differenza tra il valore reale e quello sintetico nettamente minore. In compenso, come detto precedentemente, il valore della metrica BRISQUE risulta essere più veritiero, in quanto permette di tenere conto delle percezioni soggettive.

Immagini	BRISQUE	NIQE
Sintetiche (30 epoche)	21.23 \pm 17.51	4.66 \pm 1.73
Sintetiche (40 epoche)	23.16 \pm 20.24	4.65 \pm 1.52
Sintetiche (50 epoche)	24.57 \pm 17.43	5.02 \pm 1.44
Reali	11.68 \pm 10.08	3.40 \pm 1.02

Tabella 4.4: Confronto delle metriche no-reference per il secondo allenamento: valutazione senza distinzione tra le label. I risultati sono mostrati come media \pm deviazione standard.

Valutando, invece, l'andamento differenziato per label (mostrato in tabella 4.5), possiamo notare che la label che meglio viene rappresentata è quella benigna, diversamente dal primo allenamento. Infatti si nota come la differenza tra il valore per le immagini reali e quello ottenuto per le immagini sintetizzate in cinquanta epoche è 1.64 ± 6.8 , lo stesso ragionamento non può essere effettuato per le altre label, la cui differenza risulta essere molto maggiore.

Possiamo, inoltre, notare come nel caso delle immagini sintetizzate in trenta epoche i valori risultino essere nettamente migliori per le label "Gleason 3" e "Gleason 4"; mentre per la label "Gleason 5", la condizione migliore è quella dell'allenamento con 40 epoche.

Per quanto riguarda la metrica NIQE, l'allenamento con 30 e 40 epoche, non mostra alcuna differenza e permettono di ottenere i valori migliori per tutte le label

disponibili. I dati sono mostrati graficamente in figura 4.4

Metrica	Immagini reali			
–	Ben	G13	G14	G15
BRISQUE	9.61±7.38	9.99±8.06	11.20±6.90	14.01±11.98
NIQE	2.98±0.74	3.36±0.83	3.62±0.82	3.53±1.23
Metrica	Immagini sintetiche (30 epoche)			
–	Ben	G13	G14	G15
BRISQUE	13.42±15.11	20.42±12.46	16.47±12.86	24.26±21.59
NIQE	3.65±0.98	4.74±1.05	4.71±1.08	4.82±2.29
Metrica	Immagini sintetiche (40 epoche)			
–	Ben	G13	G14	G15
BRISQUE	10.14±17.02	30.08±16.75	24.05±12.1868	19.06±21.86
NIQE	3.65±0.98	4.74±1.05	4.71±1.08	4.82±2.29
Metrica	Immagini sintetiche (50 epoche)			
–	Ben	G13	G14	G15
BRISQUE	11.25±14.12	25.27±13.69	23.02±11.11	27.24±20.23
NIQE	3.77±1.1	5.26±1.11	5.56±1.15	5.06±1.65

Tabella 4.5: Confronto delle metriche no-reference per il secondo allenamento: valutazione con distinzione tra le label. I risultati sono mostrati come media \pm deviazione standard.

Infine, sono state valutate la numerosità e l’area occupata dalle diverse label, ottenendo i valori mostrati in tabella 4.6. Confrontando i valori con quelli del primo allenamento, possiamo notare come non solo la numerosità ma anche la condizione delle aree risulti essere più bilanciata rispetto al caso precedente. La condizione della numerosità era già stata valutata numericamente in fase di allenamento, quindi il bilanciamento ottenuto era un dato previsto e ricercato.

Label	Numerosità (%)	Area (%)
Sfondo	55.89	18.59
Benigna	23.71	11.75
Gleason 3	53.63	27.83
Gleason 4	11.59	7.83
Gleason 5	77.22	15.86
Non marcato	57.05	18.13

Tabella 4.6: Valutazione percentuale dei valori medi della numerosità delle label sul numero totale di immagini e sull'area delle singole label rispetto all'area totale della patch. I valori sono ottenuti considerando il secondo allenamento effettuato.

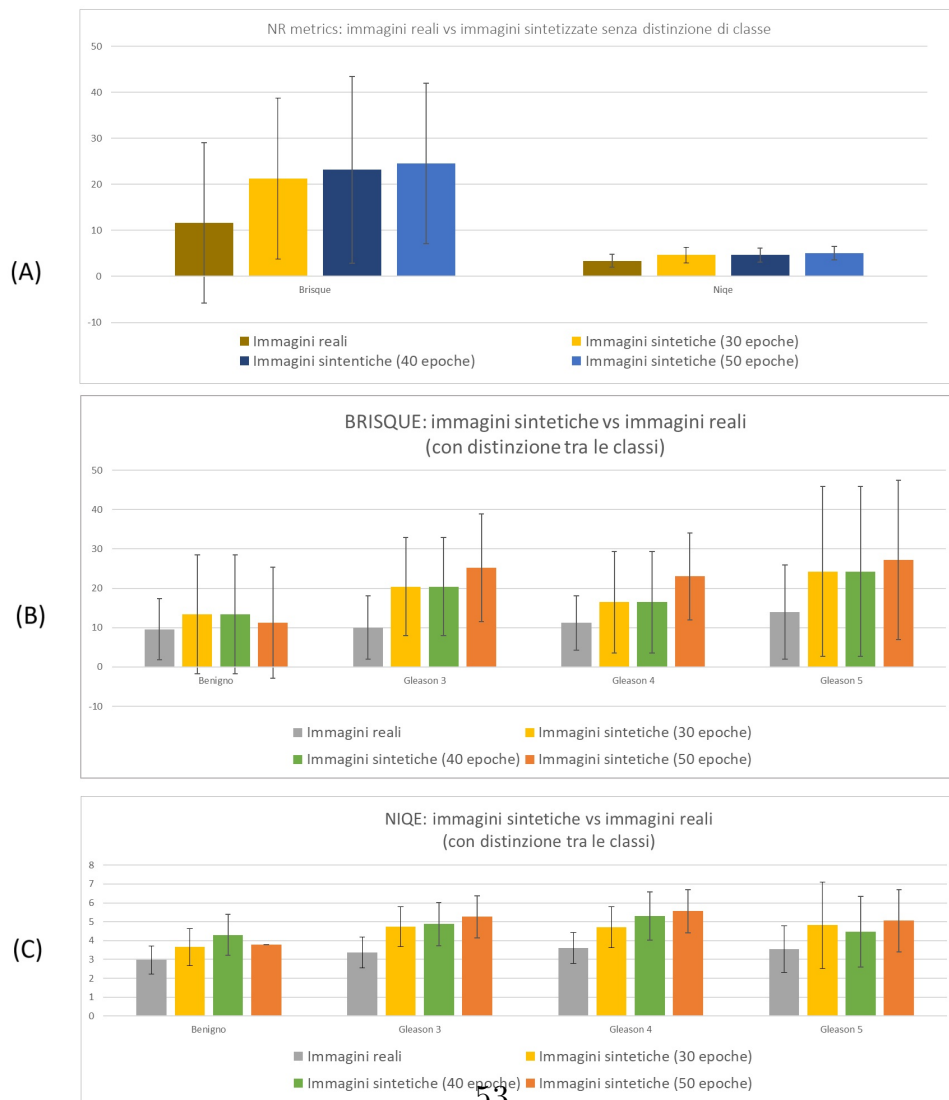


Figura 4.4: Rappresentazione grafica dei dati relativi alle NR metrics ottenuti nel secondo allenamento. In particolare: (A) andamento delle metriche senza distinzione delle label; (B) valutazione della metrica BRISQUE con distinzione tra le label; (C) valutazione della metrica NIQE con distinzione tra le label.

4.2 Texture Analysis

Infine, in questo paragrafo si vogliono analizzare i risultati dell'analisi della tessitura, i quali sono stati ottenuti solo per il secondo allenamento svolto.

Come detto nel capitolo precedente 3.8, l'analisi della tessitura permette di analizzare le caratteristiche dell'immagine andando a considerare due categorie di features, quelle del primo ordine e quelle del secondo ordine, le cui definizioni sono contenute rispettivamente nell'appendice A e nell'appendice B. L'analisi inizia con la valutazione delle features del primo ordine, ovvero quelle che sono valutate partendo dall'istogramma della scala dei grigi, che permettono di avere un'informazione "sintetica" dell'immagine, in quanto non viene tenuto conto delle informazioni relative alla spazialità. In seguito, l'analisi si concentrerà sulle features del secondo ordine, ovvero quelle ottenute partendo dalla matrice delle co-occorrenze della scala dei grigi, la quale ha tra i suoi vantaggi la possibilità di valutare le features in relazione alla spazialità che assumono i pixel rispetto all'immagine.

L'analisi è stata effettuata andando a valutare le features per le singole label e andando a valutarle al raggiungimento delle trenta epoche, delle quaranta e delle cinquanta epoche come nel caso delle metriche per lo stesso allenamento.

4.2.1 Features del primo ordine

Per semplicità di analisi, non tutte le features valutate verranno analizzate dettagliatamente, ma sono state scelte solo alcune di esse, ovvero: media, energia totale, skewness, uniformità, kurtosis, entropia e varianza. La rappresentazione grafica degli andamenti delle prime quattro features elencate è mostrata in figura 4.5, mentre per le ultime tre elencate è mostrata in figura 4.6.

Analizzando le features mostrate nella prima immagine (figura 4.5) possiamo notare come l'andamento di tutte le features non subisce delle grandi variazioni considerando l'allenamento con le trenta epoche e quello delle quaranta epoche; questo potrebbe significare che una variazione dell'andamento si può riscontrare su un numero maggiore di epoche, difatti possiamo notare come nel passaggio tra le trenta e le cinquanta epoche i vari andamenti subiscono delle differenze. Al contempo, possiamo notare come lo sfondo assuma approssimativamente lo stesso andamento nelle varie features valutate, tranne nel trend della skewness, dove insieme alla label benigna, vengono assunti dei valori negativi, diversamente dalle altre label. La presenza di un valore negativo è dovuto ad una asimmetria negativa dell'istogramma, ovvero i valori della media e della moda sono posizionati a sinistra rispetto al valore mediano. Diversamente, nel caso delle altre label siamo in presenza di un'asimmetria positiva, il che implica la presenza di media e moda alla destra del valore mediano. Per quanto concerne la valutazione dell'energia

totale delle immagini, si nota una forte congruenza dei valori reali e dell'allenamento con le cinquanta epoche nel caso delle label che mostrano la presenza tumorale. Il ragionamento, risulta essere analogo per la valutazione dell'uniformità, in cui i valori dell'allenamento con le cinquanta epoche risultano essere maggiormente prossimi a quelli reali rispetto a quelli ottenuti con gli altri allenamenti.



Figura 4.5: Andamento della media, della skewness, dell'energia totale e dell'uniformità per le cinque label (Background (sfondo), Benigna, Gleason 3, Gleason 4 e Gleason 5).

Infine, nell'analisi dell'andamento delle altre features (figura 4.6), il ragionamento avvalorato per le precedenti features, continua a ripresentarsi. Difatti i trend definiti dagli allenamenti tra le trenta e le quaranta epoche non sortiscono alcuna differenza tra loro, mentre il ragionamento non risulta essere veritiero per l'allenamento con il numero massimo di epoche. Come si può notare, per semplicità di visualizzazione, i label relativi allo sfondo sono mostrati in grafici separati a causa dei valori

discostanti che assumono rispetto alle altre label d'interesse. Da questa analisi si può notare, come nelle diverse features i valori che si ottengono, indistintamente dall'allenamento considerato, sono particolarmente discostanti rispetto al valore delle immagini reali. Possiamo, quindi, valutare come questa label risulti essere "difficile" da generare, anche se visivamente risulti essere quella di più semplice costruzione. Inoltre, un altro elemento che si può notare è come nel caso della varianza e della kurtosis l'andamento delle label sia coerente tra i vari allenamenti, con un trend ascendente nel caso della kurtosis ed uno discendente nel caso della varianza.



Figura 4.6: Andamento della kurtosis, della varianza e dell'entropia per le cinque label (Background (sfondo), Benigna, Gleason 3, Gleason 4 e Gleason 5).

4.2.2 Features del secondo ordine

Come nel caso delle features del primo ordine, anche per quelle del secondo ordine sono state analizzate nello specifico solo alcune di esse. In particolare, le features analizzate sono: l'autocorrelazione, la differenza inversa, il contrasto, la media congiunta, l'entropia differenziale, l'entropia congiunta, la varianza delle differenze (dissomiglianza) e il coefficiente di correlazione massimo (MCC).

Analizzando i grafici in figura 4.7, possiamo notare come la label dello sfondo assuma quattro differenti comportamenti:

- assume valori costanti e uguali alle immagini reali per la differenza inversa e l'entropia congiunta;
- assume valori, simili a quelli delle immagini reali, ma nettamente inferiori alle altre label nel caso di contrasto, entropia e varianza delle differenze (dissomiglianza);
- assume un trend discendente per autocorrelazione e media congiunta;
- assume un trend ascendente per il coefficiente di correlazione massimo (MCC).

Questi risultati mostrano come le differenze di questo label rispetto agli altri si mantenga anche nel momento in cui si considerano queste features, che tengono conto della spazialità e non solo del valore che assumono.

Analizzando, invece, le altre label notiamo come anche in questo caso gli allenamenti di trenta e quaranta epoche portano agli stessi risultati, diversamente dal caso dell'allenamento con le cinquanta epoche. Nel caso della label "benigna" si nota come l'allenamento ha portato ad un miglioramento dei risultati in tutte le features considerate e risulta essere un ragionamento applicabile anche alla features "Gleason 3"; questo aspetto può essere ricondotto alla maggiore semplicità di tessitura dell'immagine rispetto agli altri due gradi tumorali. Difatti, per la label "Gleason 4" e "Gleason 5" si notano degli andamenti nettamente differenti, portando anche ad avere dei valori peggiori nel caso dell'allenamento con le cinquanta epoche, come nel caso del contrasto per la label "Gleason 4". Risulta che la difficoltà venga riscontrata soprattutto in questa label, rispetto a quella "Gleason 5". Difatti possiamo notare come i valori ottenuti con un numero minore di epoche, tendano ad essere più vicini a quelli definiti dalle immagini reali. Diversamente, nel caso della label "Gleason 5" possiamo notare come il numero maggiore di epoche permetta di raggiungere dei valori più consistenti con quelli delle immagini reali.



Figura 4.7: Andamento dell'autocorrelazione, del contrasto, dell'entropia differenziale, della varianza della differenza (dissomiglianza), della differenza inversa (ID), della media congiunta, dell'energia congiunta e del coefficiente di correlazione massimo (MCC) per le cinque label (Background (sfondo), Benigna, Gleason 3, Gleason 4 e Gleason 5).

Capitolo 5

Conclusioni

In conclusione, questo lavoro di tesi aveva come scopo quello di poter portare alla luce una soluzione relativa ad una delle attuali open challenge nell'ambito dell'istopatologia digitale, ovvero la mancanza di numerosi campioni d'analisi.

Riassumendo il lavoro svolto, si è mostrato come una soluzione potrebbe essere l'utilizzo di un approccio generativo attraverso l'uso di algoritmi, come nel caso specifico l'algoritmo SPADE, che permetta di partire dal contenuto semantico delle immagini per generarne di nuove, potendo così ampliare quelli che sono gli attuali dataset relativi all'istopatologia prostatica.

L'utilizzo di questo algoritmo ha mostrato come le immagini ottenute siano coerenti a quelle reali e come questo possa essere un approccio positivo per lo scopo considerato; al contempo bisogna considerare come l'idea mostrata abbia bisogno di alcune modifiche per poter essere utilizzato nelle piene potenzialità. In particolare, possiamo valutare come le classi tumorali abbiano diverse tendenze, con una difficoltà maggiore per la generazione delle label caratterizzate da una tessitura più complessa, ovvero "Gleason 4" e "Gleason 5". Inoltre, la valutazione a diverse epoche considerate ha mostrato come all'aumentare del numero di epoche, si abbia una difficoltà maggiore relativamente alla generazione di queste label, che avevano valori migliori (soprattutto nel caso della "Gleason 4") per allenamenti con trenta o quaranta epoche; allo stesso tempo, l'aumento di sole dieci epoche non ha mostrato dei miglioramenti particolarmente evidenti, lasciando così intendere che un possibile miglioramento possa essere quello di valutare gli andamenti aumentando ulteriormente il numero di epoche, potendo così effettuare una valutazione più approfondita.

Allo stesso tempo, si è notato come la presenza del "bianco" relativamente allo sfondo possa causare delle variazioni ampie dei valori delle features, anche se la sua struttura possa sembrare così semplice. Questo potrebbe essere imputato a delle variazioni di luce non visibili ad occhio nudo, il quale tende a mostrare un elemento caratterizzato da un colore omogeneo.

Possiamo, quindi, valutare come un elemento fondamentale sia la necessità di possedere un dataset che sia bilanciato e caratterizzato dal giusto numero di elementi, per poter allenare correttamente un algoritmo di generazione semantica. Inoltre, un ulteriore elemento che sembra caratterizzare un corretto svolgimento della fase di training è relativo alla valutazione numerica delle epoche da considerare. Difatti è stato valutato come il numero delle epoche vada a intaccare l'andamento dell'analisi tissutale delle label più complesse, quali "Gleason 4" e "Gleason 5". Per concludere, si vogliono presentare diverse possibilità di studio future. La prima è quella di poter generalizzare questo algoritmo per tessuti che provengano da altri organi, che non siano quelli considerati nello studio appena descritto. Tale possibilità potrebbe essere affrontata andando a generalizzare la pipeline del lavoro. La seconda possibilità è quella di poter utilizzare un algoritmo che funga da discriminatore, quindi utilizzandolo per la valutazione primaria di una diagnosi attraverso il confronto del risultato del discriminatore con gli elementi che definiscono la classe di appartenenza. Volendo considerare quest'idea per la valutazione del grado del tessuto prostatico, alcuni elementi che possono essere considerati come fattori discriminanti sono quelli che vanno a caratterizzare i differenti stadi della patologia tumorale, ovvero l'area dei bianchi e la concentrazione cellulare all'interno dell'immagine.

Appendice A

Descrittori del primo ordine

Definizioni letterali

- *10 percentile*: valore al di sotto del quale si trova il 10% delle osservazioni;
- *90 percentile*: valore al di sotto del quale si trova il 90% delle osservazioni;
- *Energia*: misura della grandezza dei valori voxel di un'immagine;
- *Entropia*: definisce l'incertezza e la causalità nei valori dell'immagine;
- *Scarto interquartile*: differenza tra il settantacinquesimo percentile e il venticinquesimo percentile;
- *Kurtosis*: picco della distribuzione dei valori della ROI dell'immagine;
- *Massimo*: valore massimo all'interno della ROI;
- *Deviazione media assoluta*: distanza media di tutti i valori di intensità rispetto al valore medio;
- *Media*: valore medio del livello di grigio all'interno della ROI;
- *Mediana*: valore centrale nell'insieme di valori all'interno della ROI;
- *Minimo*: minimo valore all'interno della ROI;
- *Range*: estremi dell'intervallo di valori che si possono trovare all'interno della ROI;
- *Deviazione media assoluta robusta*: distanza media di tutti i valori dal valore medio calcolato sul sottoinsieme della matrice di immagini con livelli di grigio intermedi o uguali al decimo e novantesimo percentile;

- *Skewness*: definisce l'asimmetria della distribuzione dei valori rispetto al valore medio;
- *Radice della media quadratica*: media di tutti i valori di intensità al quadrato;
- *Energia totale*: valore della caratteristica energia scalato in base al volume del voxel in mm cubi;
- *Uniformità*: misura della somma dei quadrati di ciascun valore di intensità;
- *Varianza*: media delle distanze quadrate di ciascun valore di intensità rispetto al valore medio.

Definizioni matematiche

Descrittore	Definizione matematica
Energia	$\sum_{i=1}^{N_p} (X(i) + c)^2$
Entropia	$-\sum_{i=1}^{N_g} p(i) \cdot \log_2(p(i) + \epsilon)$
Kurtosis	$\frac{\frac{1}{N_p} \cdot \sum_{i=1}^{N_p} (X(i) - \bar{X})^4}{\left(\frac{1}{N_p} \cdot \sum_{i=1}^{N_p} (X(i) - \bar{X})^2\right)^2}$
Deviazione media assoluta (MAD)	$\frac{1}{N_p} \cdot \sum_{i=1}^{N_p} X(i) - \bar{X} $
Media	$\frac{1}{N_p} \cdot \sum_{i=1}^{N_p} X(i)$
Range	$\max(X) - \min(X)$
Deviazione media assoluta robusta	$\frac{1}{N_{10-90}} \cdot \sum_{i=1}^{10-90} X(i)_{10-90} - \bar{X}_{10-90} $
Radice della media quadratica	$\sqrt{\sum_{i=1}^{N_p} (X(i) + c)^2}$
Skewness	$\frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^3}{\sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2}}$
Energia totale	$V_{voxel} \cdot \sum_{i=1}^{N_p} (X(i) + c)^2$
Uniformità	$\sum_{i=1}^{N_g} p(i)^2$
Varianza	$\frac{1}{N_p} \cdot \sum_{i=1}^{N_p} (X(i) - \bar{X})^2$

Tabella A.1: Definizioni matematiche dei descrittori del primo ordine.

Appendice B

Descrittori del secondo ordine

Definizioni letterali

- *Autocorrelazione*: misura dell'entità della finezza e della grossolanità della tessitura;
- *Cluster shade (ombra del cluster)*: misura dell'asimmetria e dell'uniformità della GLCM. Una tonalità più alta dei cluster implica una maggiore asimmetria rispetto alla media;
- *Cluster prominence (prominenza del cluster)*: misura dell'asimmetria della GLCM. Un valore più alto implica una maggiore asimmetria rispetto alla media mentre un valore più basso indica un picco vicino al valore medio e una minore variazione rispetto alla media;
- *Cluster tendency (tendenza al cluster)*: misura dei raggruppamenti di voxel con valori di grigio simili;
- *Contrasto*: misura della variazione locale dell'intensità, favorendo i valori lontani dalla diagonale ($i = j$). Un valore maggiore è correlato ad una maggiore disparità nei valori di intensità tra i voxel vicini;
- *Correlazione*: valore compreso tra 0 ed 1 che mostra la dipendenza lineare dei valori del livello di grigio dai rispettivi voxel nel GLCM;
- *Differenza media*: misura della relazione tra le occorrenze di coppie con valori di intensità simili e le occorrenze di coppie con valori di intensità diversi;
- *Entropia differenziale*: misura della casualità nelle differenze dei valori di intensità del vicinato;

- *Varianza della differenza*: misura di eterogeneità che attribuisce pesi più elevati ad coppie di livelli di intensità diversi che si discostano maggiormente dalla media;
- *Differenza inversa (ID)*: misura dell'omogeneità locale di un'immagine;
- *Momento di differenza inversa (IDM)*: misura dell'omogeneità locale di un'immagine;
- *Momento di differenza inversa normalizzato (IDMN)*: misura dell'omogeneità locale di un'immagine;
- *Differenza inversa normalizzata (IDN)*: misura dell'omogeneità locale dell'immagine;
- *Misura informativa di correlazione (IMC1)*: valuta la correlazione tra le distribuzioni di probabilità. utilizzando l'informazione reciproca;
- *Misura informativa di correlazione (IMC2)*: valuta la correlazione tra le distribuzioni di probabilità;
- *Varianza inversa*: media delle distanze di ciascun valore di intensità rispetto al quadrato dei valori totali;
- *Media congiunta*: definisce l'intensità media del livello di grigio della i -esima distribuzione;
- *Energia congiunta*: misura dei modelli omogenei nell'immagine. Un'energia maggiore implica che ci sono più esempi di coppie di valori di intensità nell'immagine che si avvicinano tra loro a frequenze più alte;
- *Entropia congiunta*: misura della casualità nei valori di intensità del vicinato;
- *Coefficiente di correlazione massimo (MCC)*: misura della complessità della trama e può assumere valore compreso tra zero ed uno. Si sceglie il secondo valore massimo;
- *Massima probabilità*: indica la presenza della coppia predominante di valori di intensità vicini;
- *Somma media*: misura la relazione tra le occorrenze di coppie con valori di intensità inferiori e le occorrenze con valori di intensità più elevati;
- *Somma dell'entropia*: somma delle differenze dei valori di intensità del vicinato;
- *Somma dei quadrati*: misura nella distribuzione di coppie di livelli di intensità vicini rispetto al livello di intensità medio del GLCM.

Definizioni matematiche

Descrittore	Definizione matematica
Autocorrelazione	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \cdot i \cdot j$
Cluster prominence	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^4 \cdot p(i, j)$
Cluster shade	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^3 \cdot p(i, j)$
Cluster tendency	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^2 \cdot p(i, j)$
Contrasto	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 \cdot p(i, j)$
Correlazione	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \cdot i \cdot j - \mu_x \mu_y}{\sigma_x(i) \sigma_y(j)}$
Differenza media	$\sum_{i=1}^{N_g-1} k \cdot p_{x-y}(k)$
Entropia differenziale	$\sum_{i=1}^{N_g} p_{x-y}(k) \log_2(p_{x-y}(k) + \epsilon)$
Varianza della differenza	$\sum_{k=0}^{N_g-1} (k - DA)^2 \cdot p_{x-y}(k)$
Energia congiunta	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p(i, j))^2$
Entropia congiunta	$-\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log_2(p(i, j) + \epsilon)$
Misura informativa di correlazione	$\frac{HXY - HXY_1}{\max HX, HY} 2$
Misura informativa di correlazione	$\sqrt{1 - e^{-2(HXY_2 - HXY)}}$

Tabella B.1: Definizioni matematiche dei descrittori del secondo ordine (prima parte).

Descrittore	Definizione matematica
Momento di differenza inversa	$\sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1+k^2}$
Coefficiente di correlazione massimo	$max \left(\sqrt{\sum_{k=0}^{N_g} \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}} \right)$
Momento di differenza inversa normalizzato	$\sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + \left(\frac{k^2}{N_g^2} \right)}$
Differenza inversa	$\sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1+k}$
Differenza inversa normalizzato	$\sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + \left(\frac{k}{N_g^2} \right)}$
Varianza inversa	$\sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{k^2}$
Massima probabilità	$max(p(i, j))$
Somma media	$\sum_{k=2}^{2N_g} p_{x+y}(k) \cdot k$
Somma dell'entropia	$\sum_{k=2}^{2N_g} p_{x+y}(k) \cdot \log_2(p_{x+y}(k) + \epsilon)$
Somma dei quadrati	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 \cdot p(i, j)$

Tabella B.2: Definizioni matematiche dei descrittori del secondo ordine (seconda parte).

Bibliografia

- [1] Xiaodong Chen, Bin Zheng e Hong Liu. «Optical and Digital Microscopic Imaging Techniques and Applications in Pathology». In: *Analytical cellular pathology (Amsterdam)* 34 (gen. 2011), pp. 5–18. DOI: 10.3233/ACP-2011-0006 (cit. a p. 2).
- [2] Kai H. Hammerich, Gustavo E. Ayala e Thomas M. Wheeler. «Anatomy of the prostate gland and surgical pathology of prostate cancer». In: *Prostate Cancer*. A cura di Hedvig Hricak e PeterEditors Scardino. Contemporary Issues in Cancer Imaging. Cambridge University Press, 2008, pp. 1–14 (cit. a p. 3).
- [3] Muhammad Amin, A Khalid, Nadia Tazeen e Muhammad Yasoob. «Zonal Anatomy of Prostate». In: *Annals of King Edward Medical University* 16.3 (gen. 1), p. 138. DOI: 10.21649/akemu.v16i3.212. URL: <https://www.annalskemu.org/journal/index.php/annals/article/view/212> (cit. a p. 3).
- [4] Xudong Yu, Ruijia Liu, Lianying Song, Wenfeng Gao, Xuyun Wang e Yaosheng Zhang. «Differences in the pathogenetic characteristics of prostate cancer in the transitional and peripheral zones and the possible molecular biological mechanisms». In: *Frontiers in Oncology* 13 (2023). ISSN: 2234-943X. DOI: 10.3389/fonc.2023.1165732. URL: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1165732> (cit. alle pp. 3, 5).
- [5] John E. McNeal. «The zonal anatomy of the prostate». In: *The Prostate* 2.1 (1981), pp. 35–49. DOI: <https://doi.org/10.1002/pros.2990020105>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/pros.2990020105>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/pros.2990020105> (cit. a p. 4).
- [6] Christine H. Lee, Oluyemi Akin-Olugbade e Alexander Kirschenbaum. «Overview of Prostate Anatomy, Histology, and Pathology». English. In: *Endocrinology and Metabolism Clinics of North America* 40.3 (set. 2011), pp. 565–575. ISSN: 0889-8529. DOI: 10.1016/j.ec1.2011.05.012 (cit. a p. 4).

- [7] Anil Kapoor. «Benign prostatic hyperplasia (BPH) management in the primary care setting». In: *Can J Urol* 19.Suppl 1 (2012), pp. 10–17 (cit. a p. 6).
- [8] David Yeo e Graeme Steele. «Benign Prostatic Hypertrophy». In: *Decision Making in Medicine (Third Edition)*. A cura di Stuart B. Mushlin e Harry L. Greene. Third Edition. Philadelphia: Mosby, 2010, pp. 564–565. ISBN: 978-0-323-04107-2. DOI: <https://doi.org/10.1016/B978-0-323-04107-2.50204-0>. URL: <https://www.sciencedirect.com/science/article/pii/B9780323041072502040> (cit. a p. 6).
- [9] Mark S. Litwin e Hung-Jui Tan. «The Diagnosis and Treatment of Prostate Cancer: A Review». In: *JAMA* 317.24 (giu. 2017), pp. 2532–2542. ISSN: 0098-7484. DOI: 10.1001/jama.2017.7248. eprint: https://jamanetwork.com/journals/jama/articlepdf/2633921/jama_litwin_2017_rv_170003.pdf. URL: <https://doi.org/10.1001/jama.2017.7248> (cit. a p. 7).
- [10] Sergio Bracarda, Ottavio de Cobelli, Carlo Greco, Tommaso Prayer-Galetti, Riccardo Valdagni, Gemma Gatta, Filippo de Braud e Georg Bartsch. «Cancer of the prostate». In: *Critical reviews in oncology/hematology* 56.3 (2005), pp. 379–396 (cit. a p. 7).
- [11] Jean-Luc Descotes. «Diagnosis of prostate cancer». In: *Asian Journal of Urology* 6.2 (2019). Prostate Cancer : On the Road of Progress, pp. 129–136. ISSN: 2214-3882. DOI: <https://doi.org/10.1016/j.ajur.2018.11.007>. URL: <https://www.sciencedirect.com/science/article/pii/S2214388219300128> (cit. a p. 7).
- [12] Joseph E. Oesterling, Steven J. Jacobsen e William H. Cooner. «The Use of Age-Specific Reference Ranges for Serum Prostate Specific Anitgen in Men 60 years Old or Older». In: *The Journal of Urology* 153.4 (1995), pp. 1160–1163. ISSN: 0022-5347. DOI: [https://doi.org/10.1016/S0022-5347\(01\)67538-4](https://doi.org/10.1016/S0022-5347(01)67538-4). URL: <https://www.sciencedirect.com/science/article/pii/S0022534701675384> (cit. a p. 8).
- [13] Hedvig Hricak, Peter L Choyke, Steven C Eberhardt, Steven A Leibel e Peter T Scardino. «Imaging prostate cancer: a multidisciplinary perspective». In: *Radiology* 243.1 (2007), pp. 28–53 (cit. alle pp. 9, 10).
- [14] CJ Das, A Razik, S Sharma e S Verma. «Prostate biopsy: when and how to perform». In: *Clinical radiology* 74.11 (2019), pp. 853–864 (cit. a p. 10).
- [15] Vito Lacetera et al. «MRI/US fusion prostate biopsy: Our initial experience». In: *Archivio Italiano di Urologia e Andrologia* 88.4 (2016), pp. 296–299 (cit. a p. 10).

-
- [16] Jonathan I Epstein et al. «A contemporary prostate cancer grading system: a validated alternative to the Gleason score». In: *European urology* 69.3 (2016), pp. 428–435 (cit. a p. 11).
- [17] Peter A Humphrey, Holger Moch, Antonio L Cubilla, Thomas M Ulbright e Victor E Reuter. «The 2016 WHO classification of tumours of the urinary system and male genital organs—part B: prostate and bladder tumours». In: *European urology* 70.1 (2016), pp. 106–119 (cit. a p. 12).
- [18] Eugene U Thiessen. «Concerning a familial association between breast cancer and both prostatic and uterine malignancies». In: *Cancer* 34.4 (1974), pp. 1102–1107 (cit. a p. 13).
- [19] Famke Aeffner et al. «Introduction to digital image analysis in whole-slide imaging: a white paper from the digital pathology association». In: *Journal of pathology informatics* 10.1 (2019), p. 9 (cit. a p. 15).
- [20] Liron Pantanowitz. «Digital images and the future of digital pathology». In: *Journal of pathology informatics* 1 (2010) (cit. a p. 15).
- [21] Mark D Zarella, Douglas Bowman, Famke Aeffner, Navid Farahani, Albert Xthona, Syeda Fatima Absar, Anil Parwani, Marilyn Bui e Douglas J Hartman. «A practical guide to whole slide imaging: a white paper from the digital pathology association». In: *Archives of pathology & laboratory medicine* 143.2 (2019), pp. 222–234 (cit. a p. 16).
- [22] Byron Smith, Meyke Hermsen, Elizabeth Lesser, Deepak Ravichandar e Walter Kremers. «Developing image analysis pipelines of whole-slide images: Pre-and post-processing». In: *Journal of Clinical and Translational Science* 5.1 (2021), e38 (cit. alle pp. 16–19).
- [23] Marc Macenko, Marc Niethammer, James S Marron, David Borland, John T Woosley, Xiaojun Guan, Charles Schmitt e Nancy E Thomas. «A method for normalizing histology slides for quantitative analysis». In: *2009 IEEE international symposium on biomedical imaging: from nano to macro*. IEEE. 2009, pp. 1107–1110 (cit. a p. 17).
- [24] Adnan Mujahid Khan, Nasir Rajpoot, Darren Treanor e Derek Magee. «A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution». In: *IEEE transactions on Biomedical Engineering* 61.6 (2014), pp. 1729–1738 (cit. alle pp. 17, 20, 22).
- [25] Jared Vicory, Heather D Couture, Nancy E Thomas, David Borland, JS Marron, John Woosley e Marc Niethammer. «Appearance normalization of histology slides». In: *Computerized Medical Imaging and Graphics* 43 (2015), pp. 89–98 (cit. a p. 17).

- [26] Yukako Yagi. «Color standardization and optimization in whole slide imaging». In: *Diagnostic pathology*. Vol. 6. Springer. 2011, pp. 1–12 (cit. a p. 19).
- [27] Erik Reinhard, Michael Adhikhmin, Bruce Gooch e Peter Shirley. «Color transfer between images». In: *IEEE Computer graphics and applications* 21.5 (2001), pp. 34–41 (cit. alle pp. 20, 21).
- [28] Babak Ehteshami Bejnordi, Geert Litjens, Nadya Timofeeva, Irene Otte-Höller, André Homeyer, Nico Karssemeijer e Jeroen AWM Van Der Laak. «Stain specific standardization of whole-slide histopathological images». In: *IEEE transactions on medical imaging* 35.2 (2015), pp. 404–415 (cit. alle pp. 20, 23).
- [29] Erik Reinhard e Tania Pouli. «Colour Spaces for Colour Transfer». In: apr. 2011, pp. 1–15. ISBN: 978-3-642-20403-6. DOI: 10.1007/978-3-642-20404-3_1 (cit. a p. 20).
- [30] Matthew G Hanna, Anil Parwani e Sahussapont Joseph Sirintrapun. «Whole slide imaging: technology and applications». In: *Advances in Anatomic Pathology* 27.4 (2020), pp. 251–259 (cit. a p. 24).
- [31] Stuart C Shapiro. «Encyclopedia of artificial intelligence second edition». In: *New Jersey: A Wiley Interscience Publication* (1992) (cit. a p. 25).
- [32] Iqbal H Sarker. «Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions». In: *SN Computer Science* 2.6 (2021), p. 420 (cit. a p. 25).
- [33] Christian Janiesch, Patrick Zschech e Kai Heinrich. «Machine learning and deep learning». In: *Electronic Markets* 31.3 (2021), pp. 685–695 (cit. a p. 25).
- [34] Miao Cui e David Y Zhang. «Artificial intelligence and computational pathology». In: *Laboratory Investigation* 101.4 (2021), pp. 412–422 (cit. alle pp. 27, 34).
- [35] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville e Yoshua Bengio. «Generative adversarial networks». In: *Communications of the ACM* 63.11 (2020), pp. 139–144 (cit. a p. 28).
- [36] Alankrita Aggarwal, Mamta Mittal e Gopi Battineni. «Generative adversarial network: An overview of theory and applications». In: *International Journal of Information Management Data Insights* 1.1 (2021), p. 100004 (cit. a p. 28).
- [37] Liang Gonog e Yimin Zhou. «A review: generative adversarial networks». In: *2019 14th IEEE conference on industrial electronics and applications (ICIEA)*. IEEE. 2019, pp. 505–510 (cit. a p. 30).

-
- [38] Jing Zhang, Minhao Shao, Lulu Yu e Yunsong Li. «Image super-resolution reconstruction based on sparse representation and deep learning». In: *Signal Processing: Image Communication* 87 (2020), p. 115925 (cit. a p. 30).
- [39] Xinying Han, Yang Wu e Rui Wan. «A Method for Style Transfer from Artistic Images Based on Depth Extraction Generative Adversarial Network». In: *Applied Sciences* 13.2 (2023), p. 867 (cit. a p. 30).
- [40] Taesung Park, Ming-Yu Liu, Ting-Chun Wang e Jun-Yan Zhu. «Semantic image synthesis with spatially-adaptive normalization». In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019, pp. 2337–2346 (cit. alle pp. 30, 32).
- [41] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou e Alexei A Efros. «Image-to-image translation with conditional adversarial networks». In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1125–1134 (cit. a p. 31).
- [42] Fanzhou Wang, Chenxing Wang e Qingze Guan. «Single-shot fringe projection profilometry based on deep learning and computer graphics». In: *Optics Express* 29.6 (2021), pp. 8024–8040 (cit. a p. 31).
- [43] David Morrison, David Harris-Birtill e Peter D Caie. «Generative deep learning in digital pathology workflows». In: *The American Journal of Pathology* 191.10 (2021), pp. 1717–1723 (cit. alle pp. 33, 34).
- [44] Andrew Janowczyk e Anant Madabhushi. «Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases». In: *Journal of pathology informatics* 7.1 (2016), p. 29 (cit. a p. 35).
- [45] Muhammad Shahid, Andreas Rossholm, Benny Lövfström e Hans-Jürgen Zepernick. «No-reference image and video quality assessment: a classification and review of recent approaches». In: *EURASIP Journal on image and Video Processing* 2014 (2014), pp. 1–32 (cit. a p. 44).
- [46] Anish Mittal, Anush Krishna Moorthy e Alan Conrad Bovik. «No-reference image quality assessment in the spatial domain». In: *IEEE Transactions on image processing* 21.12 (2012), pp. 4695–4708 (cit. a p. 44).
- [47] Anish Mittal, Rajiv Soundararajan e Alan C Bovik. «Making a “completely blind” image quality analyzer». In: *IEEE Signal processing letters* 20.3 (2012), pp. 209–212 (cit. a p. 44).

Ringraziamenti

Arrivati al termine di questo lavoro, vorrei ringraziare il prof. Salvi e l'ing. Marzola per avermi permesso di poter lavorare con loro a questo progetto di tesi, lasciando che mi innamorassi ulteriormente di questo argomento.

Se sono arrivata a questo punto è grazie a mia madre, che con il suo sostegno mi ha permesso di rialzarmi e di continuare questo cammino, anche quando io stessa non riuscivo a credere in me e nelle mie capacità. Grazie mamma, senza di te non sarei arrivata a questo punto.

Un ringraziamento va a mio fratello Paolo, caposaldo della mia vita anche con i chilometri a dividerci, e alla mia famiglia, che mi ha sempre sostenuta vezzeggiandomi con il termine "ingegnere". Gli sarò eternamente grata, soprattutto a mia zia Marilena per la quale sarò sempre la sua bambina.

Un ringraziamento speciale va ai miei nonni, che avrei voluto avere qui con me per questo traguardo, ma che per uno sfortunato gioco del destino mi guardano da lontano.

Questi anni sarebbero stati difficili senza la famiglia torinese che mi ha accompagnata lungo il cammino.

Ad Asia e Serena, che saranno sempre le coinquiline con cui ho diviso cene, pianti e risate anche quando il tetto sotto cui si era non era più lo stesso per tutte.

Alla gang, che tra notate in aula studio e serate mondane, mi hanno fatto capire quanto anche il divertimento voglia la sua parte e posso dire che con loro il divertimento non manca mai. Per questo, per tutti i sorrisi e le risate, probabilmente non smetterò mai di ringraziarvi. Vorrei ringraziare soprattutto Alessandro, Sebastiano e Paola per essermi stati vicini, soprattutto nei momenti più bui.

Ad Alessia che è entrata nella mia vita per mezzo di quel galeotto esame, che ci ha fatto passare le pene dell'inferno ma che, come per Paolo e Francesca, ci ha fatto incontrare e mai lasciare.

A Giulia e Daniele, i futuri letterati, che mi hanno accolta tra loro e mi hanno aiutata a superare quest'ultimo scoglio tra giornate rinchiusi in biblioteca e pause caffè più lunghe delle ore di studio.

Infine, vorrei ringraziare gli amici di sempre. Quelli che mi hanno conosciuta al liceo con i capelli rossi e una linea di eyeliner troppo spessa. Quelli con cui adesso

sto per affrontare le responsabilità del mondo degli adulti, sempre con lo spirito di quando ci siamo conosciuti.

In particolare vorrei ringraziare Giulia che, da una vita, è sempre lì al mio fianco quando ho bisogno di un supporto o di un consiglio.

Grazie a tutti.