



**UNIVERSITÀ
DEGLI STUDI
DI PADOVA**



DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA BIOMEDICA

“Radiomica: una tecnologia moderna per il supporto alla decisione clinica”

Relatore: Prof. Giovanni Sparacino

Laureando: Enrico Bazzacco

**ANNO ACCADEMICO 2022 – 2023
Data di laurea 28/09/2023**

INDICE

ABSTRACT	4
1. INTRODUZIONE	5
1.1 Le immagini in medicina	5
1.1.1 Breve storia delle bioimmagini	5
1.1.2 Le tecnologie attuali	6
1.1.3 La diagnosi assistita da computer	6
1.1.4 Le tipologie di CAD	7
1.2 Diagnostica per immagini in oncologia	8
1.2.1 Imaging oncologico	8
1.2.2 Aspetti di fisiopatologia dei tumori rilevanti per la cura	9
1.2.3 I biomarcatori	10
1.3 L'analisi radiomica	11
1.3.1 Definizione generale	11
1.3.2 Radiomica in oncologia	11
2. IL FLUSSO RADIOMICO	13
2.1 I passi del Radiomic Flow	13
2.2 Passo 1: Acquisizione, pre-processing e filtraggio	13
2.3 Passo 2: Identificazione della ROI e segmentazione	14
2.4 Passo 3: La feature-extraction	15
2.4.1 Definizione di features di un'immagine medica	15
2.4.2 Classificazione delle textures	15
2.4.3 Textures del secondo ordine	16
2.4.4 Textures di ordini superiori	17
2.5 Passo 4: Analisi dei dati	18
2.5.1 Il problema degli high-dimensional data	18
2.5.2 Metodi statistici di dimensionality-reduction e analisi	18
2.6 Intelligenza artificiale in radiomica	19
2.6.1 Aspetti generali	19
2.6.2 Reti Neurali e Deep Learning	21
3. APPLICAZIONI DI RADIOMICA IN ONCOLOGIA	22
3.1 Diagnosi e trattamento del cancro al seno	22

3.1.1 Rilevanza della patologia e metodi di diagnosi.....	22
3.1.2 Radiomica per la previsione del sottotipo cellulare.....	23
3.1.3 Diagnosi differenziale di malignità da mammografia ed ecografia.....	24
3.2 Diagnosi e prognosi del cancro ai polmoni	26
3.2.1 Aspetti clinici	26
3.2.2 Assegnazione di un radiomic-score per valutare il grado metastatico del NSCLC	27
4. LIMITI E PROSPETTIVE DELLA RADIOMICA	30
4.1 I limiti dell'analisi radiomica	30
4.1.1 Situazione attuale della tecnologia.....	30
4.1.2 Variabilità dei risultati	31
4.1.3 Standardizzazione e strutturazione dei dati	35
4.2 Prospettive future.....	36
4.2.1 Direzioni della ricerca	36
4.2.2 Integrazione con altri dati clinici.....	36
4.2.3 Uso del deep learning.....	39
5. CONCLUSIONI	44
BIBLIOGRAFIA	46

ABSTRACT

In medicina, con il termine radiomica si intende una tecnica di analisi delle immagini di diagnostica che consente di ottenere informazioni di tipo quantitativo attraverso l'applicazione di specifici algoritmi e metodi matematici. Tali informazioni non sono generalmente rilevabili direttamente dal personale medico e fanno di questo strumento una risorsa fondamentale per il supporto alla decisione clinica. I metodi di analisi appartengono di norma all'informatica e all'analisi dei dati tradizionali, ma, in seguito al suo recente sviluppo, anche al settore dell'intelligenza artificiale.

Il campo medico dove la radiomica ha trovato maggiore applicazione è l'oncologia, dove viene usata per sviluppare diagnosi e prognosi più precise e trattamenti più mirati ed efficaci.

In questa tesi ci si propone di fornire un'introduzione alla radiomica, inserendola nel contesto della medicina oncologica tradizionale, e di analizzare poi i suddetti principi matematici su cui si fonda.

Nel Capitolo 1 viene innanzitutto presentata una breve storia del ruolo delle immagini e dei sistemi informatici in medicina, in particolare nella cura dei tumori, così da poter apprezzare i vantaggi e le potenzialità che la radiomica porta con sé. Nel Capitolo 2 viene esaminato, in ogni suo step, il processo di analisi proprio della tecnologia, ponendo attenzione sulle peculiarità e difficoltà che si pongono davanti ad ogni fase che lo compone, il tutto senza entrare troppo nei dettagli tecnici. In seguito, nel Capitolo 3, vengono presentati alcuni casi studio rappresentativi dei risultati ottenuti grazie all'uso di questa tecnica nella diagnosi e prognosi del cancro. Infine, nel Capitolo 4 si analizzano le attuali limitazioni e sfide insieme alle frontiere della ricerca in radiomica. Viene riportato un quadro generale della letteratura sull'argomento, con cui si mettono in evidenza i principali limiti riscontrati all'interno degli studi, le direzioni in cui la ricerca intende proseguire e le iniziative che sono nate negli ultimi anni per sostenerne lo sviluppo.

1. INTRODUZIONE

1.1 Le immagini in medicina

1.1.1 Breve storia delle bioimmagini

L'avvento delle immagini in medicina risale al 1895 quando Wilhelm C. Röntgen scopre i raggi X ed effettua la prima proiezione radiografica di un tessuto umano per mezzo di un tubo catodico. Da quel momento in poi si apre una strada nell'indagine clinica che rivoluzionerà la medicina. Sino ad allora, infatti, osservare i tessuti umani interni era possibile solamente con metodi invasivi e spesso praticabili solo post mortem. Con la scoperta dei raggi X si rende possibile visualizzare, su opportuni supporti fisici, tessuti interni senza ledere, o almeno in modo non percepibile (al tempo non si conoscevano i pericoli delle radiazioni ionizzanti), il soggetto osservato [1]. Nei primi anni del '900 viene sviluppata, in campo navale, la tecnologia SONAR e nel 1942 si ha la prima applicazione degli ultrasuoni in medicina. Negli anni '40 si eseguono le prime tomografie, ottenute facendo ruotare il tubo radiogeno attorno al paziente, ma bisogna aspettare più di mezzo secolo e lo sviluppo di mezzi informatici sufficientemente potenti per riuscire a ricostruire un'immagine tomografica a partire da semplici proiezioni radiografiche. Nel 1971 viene eseguita la prima TAC con uno scanner a tomografia computerizzata progettato da Sir Godfrey Hounsfield, il quale ricevette il premio Nobel 8 anni più tardi per tale invenzione.

Contemporaneamente, a partire dagli anni '50, si sviluppano anche le prime tecniche di diagnostica basate sull'utilizzo dei radionuclidi, ovvero elementi in grado di emettere radiazioni grazie al fenomeno del decadimento radioattivo. Con l'invenzione della gamma-camera e il suo uso combinato con la tomografia computerizzata nasce la medicina nucleare, con gli esami PET e SPECT. Negli stessi anni viene scoperto il fenomeno della risonanza magnetica e successivamente viene ideata l'MRI, applicata per la prima volta al campo medico nella discriminazione di tessuti [2]. Nel 1990 viene scoperto l'effetto BOLD (Blood Oxygenation Level Dependent) che apre la strada all'imaging funzionale con il quale è possibile osservare il funzionamento di organi e tessuti. Da qui in poi iniziano i primi tentativi di combinare più tecniche diagnostiche al fine di ottenere immagini con un contenuto di maggior valore clinico [3].

1.1.2 Le tecnologie attuali

Tutte le tecniche di imaging sopramenzionate hanno, nel tempo, subito mutamenti ed evoluzioni volti a migliorare la qualità delle immagini prodotte eliminando rumori ed artefatti e massimizzando le risoluzioni spaziale e temporale. Allo stesso tempo si è cercato di minimizzare gli effetti nocivi che l'uso di tali tecniche ha sull'organismo umano (in particolare quelle con radiazioni ionizzanti, quindi raggi X e gamma), riducendo i tempi di esposizione e l'intensità delle radiazioni emesse, giungendo ad efficaci compromessi in termini di sicurezza e impatto sulla salute [3]. Ancora oggi i vari metodi sono estensivamente utilizzati partendo dalla radiografia e l'ecografia ad ultrasuoni per gli esami più semplici e di routine sino a tecniche di combinazione di PET-MRI, PET-CT e fMRI per l'imaging funzionale ed esami più complessi.

1.1.3 La diagnosi assistita da computer

Nel XXI secolo le innovazioni tecnologiche, affiancate al miglioramento delle prestazioni dei computer, hanno acconsentito ad un sempre più massiccio utilizzo dell'informatica in medicina con l'obiettivo di migliorare l'accuratezza delle diagnosi, l'efficacia di trattamenti e terapie ed anche l'efficienza dei sistemi sanitari. Prima ancora della nascita della ricerca sulle immagini mediche, diversi ospedali in Europa e USA possedevano già degli archivi dove poter conservare e gestire le immagini di diagnostica chiamati PACS (Picture Archiving and Communication System). A partire dagli anni '80 si iniziò quindi ad ipotizzare l'uso dei computer per l'assistenza alla diagnosi basata su immagini e l'esistenza dei PACS divenne fondamentale a questo scopo. Inizialmente gli sforzi dei ricercatori si concentrarono sullo sviluppo di sistemi di *Automated Computer Diagnosis*, ovvero software informatici capaci di svolgere in maniera automatica le diagnosi a partire da immagini ed esiti di altri esami [4]. Si riteneva che in futuro questi sistemi avrebbero sostituito completamente la figura del radiologo, ma, a causa della scarsa potenza dei computer dell'epoca e l'assenza di efficienti algoritmi di elaborazione, i risultati ottenuti furono poco soddisfacenti. Dopo quasi 30 anni di ricerca si comprese il limite di tale tecnologia e si diffuse un pensiero alternativo secondo cui i computer sarebbero stati più utili se usati come supporto alla decisione clinica ed al lavoro del medico radiologo [5]. Nacquero così i CADs (Computer-Aided Diagnosis systems), degli strumenti in grado di cambiare radicalmente la qualità dei servizi offerti dai sistemi sanitari. L'obiettivo di un CAD è quello di integrare dati di diverso tipo ottenuti dagli esami ospedalieri, anche in diversi momenti temporali, in modo da migliorare la diagnosi di patologie e disturbi. Al giorno d'oggi la mole di dati clinici è di una dimensione tale che per il personale medico diventa difficile tenere in considerazione tutti i molteplici fattori di diversa natura esistenti al momento di

formulare una diagnosi e predisporre una terapia. Risulta perciò di estrema utilità disporre di sistemi in grado di analizzare i dati di un singolo paziente e metterli in relazione con quelle di altri pazienti attraverso potenti metodi statistici per fornire un sostegno ai medici.

Un altro problema che l'uso dei CAD in medicina aiuta a risolvere è quello della soggettività e variabilità delle diagnosi effettuate da operatori umani. Non è infrequente che medici diversi abbiano opinioni differenti sullo stesso caso clinico e fattori come l'esperienza ed il contesto (stress, stato d'animo, ecc.) influenzano fortemente la qualità delle diagnosi fornite dal personale specializzato. Un'analisi operata da un computer permette, inoltre, di fornire delle misure e dei riferimenti oggettivi e precisi che descrivano quantitativamente la patologia che si sta indagando permettendo di costruire metodi di classificazione e scale di valutazione della gravità o dello stadio di avanzamento della stessa [6].

1.1.4 Le tipologie di CAD

I primi CAD erano basati sugli *expert systems*, sistemi dove l'utilizzatore interagisce con un motore inferenziale di tipo induttivo o deduttivo applicato alla *knowledge base*, un insieme di fatti oggettivi e regole che il motore usa per compiere le inferenze. Questa base di partenza è predisposta da un operatore esterno e da essa un algoritmo basato su enunciati di tipo *if-then* produce, attraverso passaggi logici, un'inferenza conclusiva, imitando il ragionamento umano. Alcune varianti di questo sistema includono delle regole associate ad una valutazione probabilistica. I limiti principali dei sistemi esperti riguardano la possibilità di verificare se le inferenze prodotte siano effettivamente corrette e logicamente coerenti, specialmente quando il numero di regole è molto alto e risulta computazionalmente oneroso operare qualsiasi tipo di controllo [6]. Per questo motivo si è nel tempo deciso di perseguire metodi basati sull'apprendimento automatico. Gli approcci basati sul *machine learning* si dividono in supervisionati e non-supervisionati nella relazione con l'operatore: nel primo caso i dati su cui il modello esegue il *training* sono preventivamente classificati con determinati criteri, nel secondo invece sono privi di ogni informazione aggiuntiva (*raw data*). Con l'uso dell'intelligenza artificiale si superano molti dei problemi presenti nei sistemi esperti e si ottengono modelli molto più versatili e prestazionali [6].

1.2 Diagnostica per immagini in oncologia

1.2.1 Imaging oncologico

Le bioimmagini rivestono un ruolo fondamentale nel campo della medicina oncologica grazie ai loro numerosi vantaggi: permettono il monitoraggio in tempo reale, un accesso non invasivo e senza danneggiamento tissutale e inoltre funzionano su una larga scala sia spaziale che temporale (possono ritrarre porzioni estese del corpo e sono agevolmente ripetibili). Se integrate con altri strumenti diagnostici come l'analisi di fluidi e tessuti in vivo diventano fondamentali nel supporto alla decisione clinica [7].

Attualmente ogni tecnica di imaging ha un uso consolidato in medicina per specifiche patologie, di seguito viene riportato un elenco dei tipici esami di oncologia effettuati con le tecnologie più diffuse:

- I raggi X vengono utilizzati per proiezioni del torace in casi di tumore al polmone, pancreas, fegato, ecc., un uso fondamentale è quello nella mammografia radar.
- La medicina nucleare viene usata in tutte le fasi del trattamento del cancro e risulta molto utile nel monitoraggio della terapia e la valutazione dello stadio di avanzamento e del grado di malignità.
- Gli ultrasuoni sono molto utilizzati per la diagnosi di tumore alla tiroide, seno, prostata, fegato, pancreas ecc. poiché essendo innocui permettono di effettuare diversi esami a breve distanza l'uno dall'altro.
- L'MRI, oltre che per la localizzazione delle lesioni, è estremamente utile per la differenziazione dei tessuti e per analizzare e misurare la risposta alle terapie.

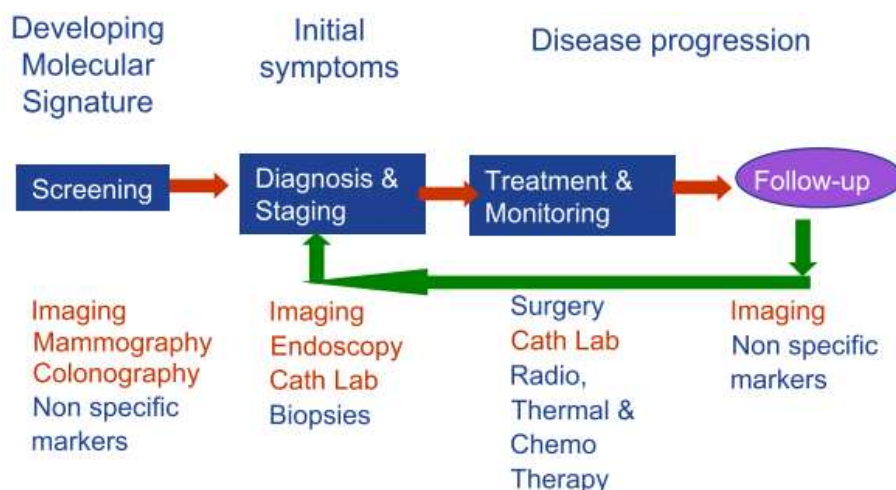


Figura 1: Immagine che schematizza l'attuale ruolo di varie procedure mediche di oncologia nelle diverse fasi della cura. In particolare, le immagini si inseriscono nello screening, la diagnosi ed il follow-up. Immagine tratta da [7].

1.2.2 Aspetti di fisiopatologia dei tumori rilevanti per la cura

Per poter comprendere come l'analisi delle immagini di diagnostica apporti un beneficio nel trattamento dei tumori maligni è necessario considerare alcuni degli aspetti fondamentali della fisiopatologia oncologica estremamente rilevanti ai fini della cura del paziente.

È ben noto che i tumori siano patologie dinamiche e instabili che vanno incontro a diverse mutazioni nel corso del tempo. Queste mutazioni sono la causa della trasformazione di una neoplasia benigna in una maligna e costituiscono l'ostacolo più importante nella disposizione di terapie, essendo alla base della perdita di efficacia dei farmaci nel tempo.

La locuzione *tumor heterogeneity* (eterogeneità tumorale) è un termine ombrello che comprende al suo interno molteplici aspetti della mutabilità di un tumore caratterizzanti la sua fisiologia e biologia.

In primo luogo, si definisce eterogeneità inter-tumorale la differente caratterizzazione cellulare di tumori dello stesso tipo istologico in pazienti diversi. Essa è da tempo conosciuta e si ritiene che sia causata principalmente da fattori paziente-specifici ed ambientali.

L'avvento della genomica ha, invece, permesso la definizione di un altro tipo di eterogeneità, detta intra-tumorale, che riguarda le cellule tumorali di un singolo paziente.

La *intratumoral heterogeneity* si distingue in spaziale e temporale. L'eterogeneità spaziale si riferisce all'esistenza di differenze genotipiche e fenotipiche tra le cellule del tumore primario e quelle delle metastasi o tra le cellule della stessa massa cancerosa. L'eterogeneità temporale riguarda le modificazioni genetiche a cui le cellule tumorali incorrono nel tempo.

Per spiegare tali attributi sono state sviluppate diverse teorie sull'evoluzione e la biologia del cancro, la più accreditata imputa l'eterogeneità di un tumore al fenomeno della *genomic instability* e spiega questa serie di mutamenti attraverso la *clonal evolution/selection hypothesis*. Tale ipotesi, formulata nel 1976 da Peter Nowell, afferma che l'iniziazione di un tumore avviene in modo stocastico, iniziando con un cambiamento in una cellula, precedentemente non maligna, che le conferisce un vantaggio selettivo di crescita e conduce alla proliferazione neoplastica. Conseguentemente, l'instabilità genomica crea un'addizionale diversità genetica nel tessuto, che, soggetto alla pressione evolutiva, conduce alla nascita di sottopopolazioni cellulari eterogenee [8].

Il problema appena descritto ha importanti implicazioni nel trattamento del cancro, specialmente nell'applicazione di terapie mirate. Si ritiene infatti che la diversità intracellulare sia la causa principale dell'insuccesso di trattamenti basati sulla classificazione molecolare e della loro progressiva perdita di efficacia nel periodo di cura [9]. Un altro aspetto da considerare è quello che le stesse terapie siano sospettate di causare un'eccessiva differenziazione delle cellule, esercitando loro stesse una selezione evolutiva, e una conseguente vanificazione

dell'effetto di farmaci specifici, designati per funzionare nel contesto patologico iniziale. Risulta pertanto necessario sviluppare metodi di indagine e monitoraggio dei tumori nel tempo capaci di identificare questi cambiamenti in modo meno invasivo e dannoso possibile.

1.2.3 I biomarcatori

Un biomarcatore, in inglese *biomarker*, è un indice dello stato di un dato sistema biologico.

In medicina un biomarcatore può essere costituito da una molecola o un composto, ad esempio antigeni, metaboliti o sostanze esogene. In oncologia il ruolo dei biomarcatori è di fondamentale importanza. Essi, infatti, forniscono informazioni riguardo la presenza di specifici antigeni, proteine o geni, utili nella caratterizzazione e descrizione del tumore. Sulla base di queste informazioni vengono poi decisi i farmaci da usare per la chemioterapia e l'immunoterapia mirate. Gli stessi indici forniscono poi informazioni sull'evoluzione della patologia e quindi l'efficacia dei trattamenti adoperati.

Esistono diverse tipologie di biomarcatori ottenuti con metodi che possono essere invasivi, come biopsie, micro-invasivi, come quelli basati sulle analisi del sangue, o non invasivi come quelli ottenuti da immagini. Un fattore che porta a preferire metodi di indagine dei marcatori tumorali meno invasivi possibile è l'eterogeneità tumorale che, come spiegato prima, implica rilevanti alterazioni nell'espressione delle molecole per mirare la terapia, rendendo necessario eseguire esami e analisi con elevata frequenza temporale e su un'ampia scala spaziale (più siti di ispezione) [9]. È infatti stato osservato che il successo di un esame istologico nel caratterizzare adeguatamente i sottotipi cellulari di una massa tumorale dipende proporzionalmente dal numero di campioni di tessuto in siti diversi che sono stati prelevati.

Il principale metodo di indagine di biomarcatori non invasivo sfrutta il sangue per ricercare specifiche molecole nel plasma e, misurandone la concentrazione, individuare i casi sospetti di neoplasie. Una tecnica più avanzata prevede l'analisi del ctDNA (cell-free tumor DNA), ovvero frammenti di DNA proveniente da cellule tumorali, da cui è possibile anche determinare il tipo di cancro [8].

1.3 L'analisi radiomica

1.3.1 Definizione generale

L'analisi radiomica consiste nell'estrazione di dati numerici, dalle immagini di diagnostica, al fine di elaborarli ed ottenere informazioni aggiuntive sulla patologia studiata. Il principio alla base di questa tecnologia è l'idea che le immagini contengano al loro interno informazioni sulla fisiologia e la biologia non rilevabili dalla semplice osservazione da parte di un operatore.

Il termine segue la costruzione lessicale delle discipline omiche, quelle branche della scienza, in particolare della biologia molecolare, che a partire da grandi moli di dati traggono informazioni attraverso complessi procedimenti analitici. Tra le principali si ricordano la Genomica (Genoma - geni), la Proteomica (Proteoma - proteine) e la Trascrittomica (Trascrittoma - RNA messaggero). Radiomica, quindi, altro non è che l'applicazione del suffisso *-omica*, con il significato di studio su larga scala, al termine radio, rappresentativo delle immagini biomediche.

1.3.2 Radiomica in oncologia

Il campo di applicazione dove questa tecnica ha riscontrato il maggior successo nei risultati ottenuti è proprio quello oncologico.

In primo luogo, l'analisi radiomica consente di integrare informazioni sulla localizzazione e le dimensioni della lesione osservata, costituendo un'evoluzione naturale dei CAD.

Le potenzialità di questa nuova tecnologia nel trattamento del cancro sono ben espresse dalla cosiddetta ipotesi radiomica [10], con cui si esprime l'idea fondamentale alla base del suo sviluppo: le immagini di diagnostica, profondamente, riflettono le caratteristiche genomiche e proteomiche della patologia tumorale rendendole caratterizzabili quantitativamente attraverso opportuni metodi di analisi.

Concretamente, quindi, la radiomica consente di descrivere il fenotipo del tumore e aspetti della sua fisiologia, di stratificare le lesioni e infine di monitorare facilmente l'andamento delle terapie e il follow-up del trattamento, alla pari di altri metodi di indagine oggi consolidati.

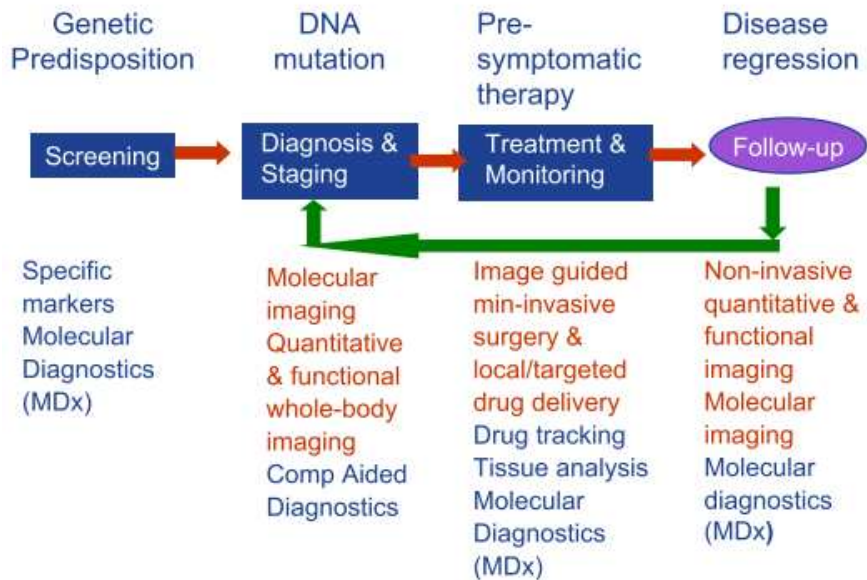


Figura 2: Evoluzione della figura 1 in cui si schematizza il ruolo che le immagini possono ricoprire in oncologia grazie alla radiomica. Qui le immagini diventano essenziali strumenti, oltre che nella diagnosi e follow-up, anche nello sviluppo del trattamento e nel monitoraggio della cura. Immagine tratta da [7].

In particolare, questo tipo di analisi, date le presunte potenzialità di inferenza genomica, si presta bene alla definizione di biomarcatori a partire dalle immagini. Allo scopo, la radiomica ha già trovato applicazione in neuro-oncologia, partendo da immagini di MRI [11], e nel trattamento del cancro dell'orofaringe (OPSCC) [12], con l'identificazione di caratteristiche nelle immagini di CT legate a marcatori sanguigni.

Per comprendere effettivamente come questo tipo di analisi consenta di ottenere risultati utili nella diagnosi dei tumori risulta necessario analizzarne il processo di implementazione. Nel prossimo capitolo, questo processo, comunemente costituito da alcuni passaggi ed una struttura consolidati e perciò denominato flusso radiomico, sarà esposto in ogni sua fase presentandone anche gli aspetti più critici per il suo utilizzo in medicina.

2. IL FLUSSO RADIOMICO

2.1 I passi del Radiomic Flow

L'analisi radiomica è un processo complessivamente semplice che si compone di una serie di step tra loro separati in quello che viene definito il Radiomic Flow. Questo flusso di lavoro è standardizzato nella costituzione, ma ogni singola fase può essere condotta in modi diversi che portano con sé sfide e problematiche specifiche [13].

Lo schema generale di flusso radiomico prevede:

1. acquisizione e filtraggio dell'immagine;
2. definizione della regione di interesse e selezione della porzione da analizzare (detta segmentazione);
3. estrazione di quantificatori numerici dell'immagine;
4. analisi dei dati estratti per ricavare informazioni di significato clinico.

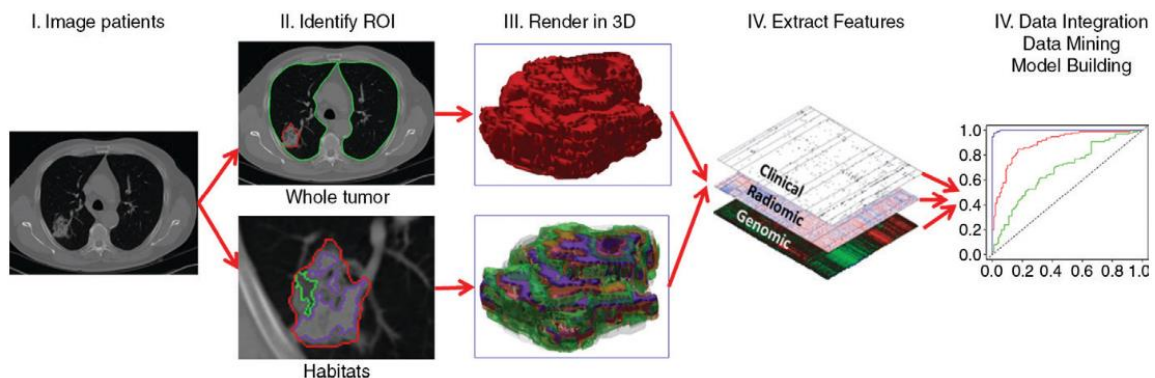


Figura 3: Schema del Radiomic Flow. I punti II e III corrispondono alla fase 2 dell'elenco presentato sopra. Immagine tratta da [13].

Nei paragrafi che seguono verranno presentate nel dettaglio tutte le fasi del workflow, discutendone limiti e difficoltà di attuazione, ma anche il ruolo che ricoprono come tassello del puzzle radiomico.

2.2 Passo 1: Acquisizione, pre-processing e filtraggio

Solitamente le immagini che vengono usate per l'analisi radiomica sono direttamente ottenute dal sistema di acquisizione, senza ulteriori elaborazioni. Alcune possibili operazioni di pre-processing effettuate sono lo *smoothing* con media mobile, l'applicazione di filtri Gaussiani per la riduzione del rumore e tecniche di *enhancement* tramite equalizzazione dell'istogramma,

deblurring, e ridimensionamento. Alcuni studi hanno, inoltre, dimostrato la dipendenza dei risultati dai valori di discretizzazione e la risoluzione dei *bin* dell'immagine [14].

2.3 Passo 2: Identificazione della ROI e segmentazione

L'Identificazione dei volumi tissutali di valore prognostico (*Region of Interest*, ROI) è il fulcro della pratica radiologica in oncologia. Al momento della diagnosi il cancro può essere rilevato in un sito tumorale singolo o in siti multipli, dato che la maggior parte dei pazienti con metastasi tumorali ha lesioni multiple. In entrambi gli scenari è necessario identificare i tumori certi e quelli sospetti con opportuni volumi di interesse. In radiomica, sotto-volumi di interesse possono essere compresi nel dataset di analisi per aumentare le informazioni di partenza per l'analisi: la filosofia di base è, infatti, di acquisire quanti più dati possibile in ingresso e operare il mining (l'estrazione di informazione da tali dati) a valle del processo. Questa scelta metodologica è guidata dalla consapevolezza che filtrare i dati in ingresso sarebbe inefficiente e presupporrebbe conoscenze a priori riguardanti la loro rilevanza ai fini della costruzione di sistemi di supporto alla diagnosi e prognosi [13].

La segmentazione è la parte più critica e che pone più sfide di tutto il workflow radiomico, poiché definisce il volume di dati su cui verrà fatta l'analisi e risulta di difficile esecuzione.

I tumori hanno, infatti, bordi irregolari difficili da delineare, motivo per cui vi è ancora un aperto dibattito sul tipo di metodo da utilizzare per identificarli. Il consenso generale è che il modo più efficace sia quello di usare software per la segmentazione automatica con la supervisione di un operatore che si occupa di rifinire il lavoro svolto dal computer [13].

La segmentazione manuale risulta di uso comune solamente per la delineazione del volume per la radioterapia, ma si dimostra di scarsa solidità nel campo delle analisi avanzate: piccole variazioni in questa operazione comportano grandi differenze nei risultati finali.

I sistemi semi-automatici hanno dimostrato successo nell'eliminare tali problemi di *consistency* e aumentare l'efficienza del processo e risultano perciò di maggiore fruizione soprattutto nel campo oncologico, dove l'utilizzo di algoritmi particolarmente avanzati permette di ridurre al minimo l'intervento umano. In altri campi della radiomica, invece, la segmentazione automatica, consentendo di distinguere con precisione tessuti come quello osseo o organi quali i polmoni, garantisce risultati soddisfacenti in quanto a qualità e coerenza [14].

2.4 Passo 3: La feature-extraction

2.4.1 Definizione di features di un'immagine medica

Le immagini di medicina sono immagini in *grayscale*, perciò sono digitalmente rappresentate da una matrice di pixels, o *voxels* nel caso tri-dimensionale, scalari contenenti un livello di grigio (i concetti espressi da qui in avanti per i pixels saranno naturalmente estendibili ai voxels). Le *features*, letteralmente caratteristiche, di un'immagine sono quei dati estraibili da essa che permettono di descriverla quantitativamente. La loro estrazione costituisce il passaggio centrale di tutta l'analisi radiomica.

Sulla base del tipo di informazione che conferiscono, le features possono essere distinte in semantiche ed agnostiche: le prime descrivono e quantificano caratteristiche del tumore, nel caso oncologico, generalmente riconosciute nella pratica clinica, quelle agnostiche, invece, forniscono misure ed informazioni sull'eterogeneità tumorale, ma non hanno una corrispondenza esatta nel vocabolario medico. Più precisamente queste ultime non si riferiscono a caratteristiche definite a priori o a fenomeni conosciuti e legati da rapporti di causa-effetto, ma descrivono rapporti di correlazione tra aspetti concreti della patologia [13]. Per questo motivo esse vengono chiamate *texture* features, dove con textures si intendono le strutture spaziali dell'intensità del livello di grigio che si ripetono periodicamente nell'immagine o in una sua regione [15]; di seguito, come avviene in letteratura, i due termini saranno usati come sinonimi. Nella cura del cancro, questa fase è la concretizzazione dell'ipotesi alla base della radiomica, ovvero che la presenza di determinate features corrisponda a precisi aspetti fenotipici e proteo-genomici della lesione interessata.

2.4.2 Classificazione delle textures

L'analisi delle textures di un'immagine digitale può essere compiuta con un'ampia varietà di metodi matematici afferenti a diversi rami della teoria dei segnali, dell'informazione e della statistica. Fra questi rientrano gli approcci classici spettrali, strutturali, probabilistici, basati sui grafi, su modelli, sull'entropia e quelli moderni basati sull'apprendimento automatico, nello specifico metodi *vocabulary-learned* e di *deep-learning* [16].

Al momento, nel campo medico e nella ricerca, si riscontra un uso preponderante dei metodi di tipo probabilistico e di *machine learning*, dovuto alla loro facile implementazione e allo sviluppo recentemente subito.

Nella pratica, le textures vengono classificate sulla base del tipo di dato che rappresentano e dalla sua dimensione (in termini statistici) in textures del primo ordine, del secondo ordine e di

ordini successivi. Le texture del primo ordine descrivono la distribuzione dei valori dei pixel senza considerare relazioni spaziali fra loro e sono calcolate a partire dall'istogramma dei valori di grigio, generalmente comprendono:

- Descrittori di range e dispersione della distribuzione come massimo/minimo, media, deviazione standard e scarto medio assoluto
- Il momento di asimmetria o *skewness* e la curtosi (indice di appiattimento)
- Descrittori di uniformità e casualità, energia ed entropia rispettivamente [14].

I vantaggi principali di tali features sono l'invarianza a rotazione e inclinazione e il basso costo computazionale, il limite è che sono in grado di fornire solo informazioni di tipo globale, senza riuscire ad isolare porzioni dell'immagine [16].

2.4.3 Textures del secondo ordine

Le textures del secondo ordine mettono in relazione il valore di grigio di un voxel con quelli a lui adiacenti, analizzando piccole regioni dell'immagine ad ogni iterazione. Esse sono, infatti, calcolate attraverso matrici che determinano quanto spesso un pixel si trova in una determinata relazione di intensità con un valore di riferimento.

Le matrici principalmente utilizzate sono:

- La Grey Level Co-Occurrence Matrix (GLCM) che considera un pixel ed il successivo e si compone contando il numero di volte in cui un dato valore di grigio è seguito da ogni altro valore

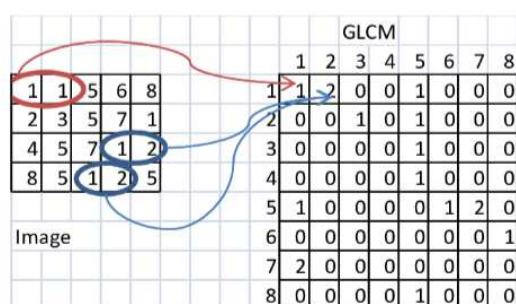


Figura 4: Schema di calcolo della GLCM (da [15]).

- La Grey Level Run-Length Matrix (GLRLM) che costruisce *run* di pixels con lo stesso livello di grigio lungo specifiche direzioni; solitamente le direzioni di calcolo sono quattro e per ognuna di esse viene costruito poi un istogramma
- La Neighbourhood Grey Level Difference Matrix (NGLDM) calcola la differenza di valore tra un pixel [voxel] e gli 8 [26] pixel [voxel] adiacenti
- La Grey Level Zone-Length Matrix (GLZLM) che è creata mettendo in un asse della matrice le differenti intensità di grigio, mentre nell'altro le possibili lunghezze delle run.

- La Grey level distance zone matrix (GLDZM)
- La Neighbouring Gray-Tone Difference Matrix (NGTDM).

Dalle matrici appena elencate si calcolano le vere e proprie textures, specifiche per ognuna. Dalla GLCM, ad esempio, si ricavano descrittori delle differenze di intensità, come contrasto e omogeneità, descrittori dell'ordine, entropia ed energia, e descrittori statistici quali massimo, media e correlazione [15]. Dalla NGLDM si ottengono descrittori di grossolanità e contrasto della regione [17].

2.4.4 Textures di ordini superiori

Le textures di ordine superiore analizzano relazioni tra i valori di grigio di insiemi di più di due pixels, per questo sono spesso più robuste al rumore dell'immagine. Tra queste si annovera l'operatore LBP (Local Binary Pattern) il quale genera un pattern binario per ogni pixel sulla base della relazione del suo livello di grigio con gli 8 adiacenti in questo modo: se il valore di grigio del pixel della *neighbourhood* (adiacenza) è maggiore di quello del centro allora esso assume valore 1, altrimenti il valore è 0, da qui si costruisce un pattern leggendo i valori in senso orario o antiorario e partendo da un pixel dipendente dal tipo di variante adottata. Il numero ottenuto può essere scritto su un'altra base per l'elaborazione successiva.

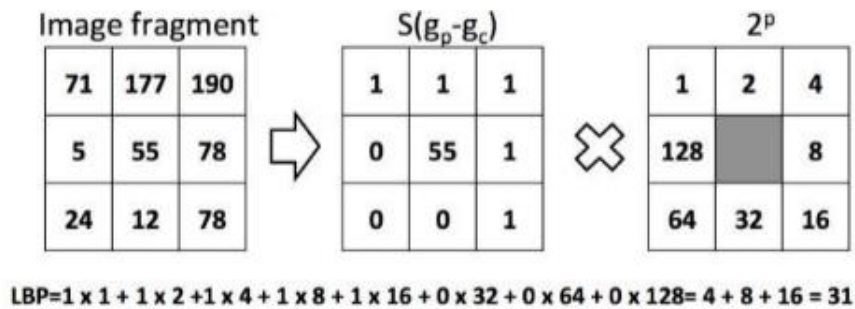


Figura 5: Esempio di calcolo del LBP (da [15]).

Talvolta si possono costruire pattern a tre valori ottenendo un operatore LTP (Local Ternary Pattern). L'LBP misura i pattern di transizione di intensità dei valori di grigio e viene usato molto spesso in versioni più complesse ed in combinazione con altri metodi, data la sua semplicità di calcolo.

Tra i metodi spettrali quello più frequentemente usato è la trasformata *Wavelet*, che lavorando sia nel dominio spaziale che in quello delle frequenze, fornisce informazioni sulle variazioni spaziali di risoluzione.

I modelli matematici frattali permettono di individuare pattern di auto-similarità in diverse scale all'interno dell'immagine mentre i funzionali di Minkowski forniscono misure relative alle

forme presenti (volume, superficie e curvatura), per le loro funzioni sono entrambi usati sovente [16].

2.5 Passo 4: Analisi dei dati

2.5.1 Il problema degli high-dimensional data

In statistica, la *dimensionality* di un dato appartenente ad un test può essere definita come il numero di variabili (features) descritte dal dato stesso, inteso come singolo campione del test. Gli *high-dimensional data* sono, quindi, dati dove la dimensionalità, corrispondente al numero di variabili osservate attraverso il dato, è un numero vicino o superiore al numero di osservazioni compiute. Questo aspetto caratterizza tutti i test radiomici che, appunto, estraggono un alto numero di descrittori e features da un numero contenuto, e spesso inferiore, di immagini di partenza. I dati di questo tipo portano con sé molteplici complicazioni nella trattazione e gestione che non si presentano nella statistica classica, caratterizzata da dati a bassa dimensione, tanto che viene usato il termine *curse of dimensionality* (letteralmente maledizione della dimensionalità) per definirli. Nonostante ciò, sono sempre più comuni in diversi ambiti della ricerca scientifica come conseguenza dell'automatizzazione della loro raccolta. I problemi principali riguardano la difficoltà di isolare variabili di risposta (o dipendenti), la presenza di forte correlazione interna al *data-set* ed il rischio di *overfitting* dei modelli costruiti. Per ovviare alle problematiche appena esposte i ricercatori ricorrono a diverse strategie, proprie di questo ramo della statistica, per la riduzione della dimensione dello spazio studiato, attraverso la selezione e classificazione delle variabili, come i metodi di regressione e *clustering*.

2.5.2 Metodi statistici di dimensionality-reduction e analisi

Una revisione della letteratura riguardante l'applicazione della radiomica in oncologia permette di individuare come metodo più diffuso, e contestualmente usato nella fase di classificazione dei dati ottenuti dalla feature extraction, il LASSO (Least Absolute Shrinkage and Selection Operator). LASSO è un metodo di analisi della regressione di un *data-set* usato per migliorare la precisione e l'interpretabilità dei modelli statistici; esso opera alterando il processo di *model-fitting* in modo che solo una porzione di tutte le variabili disponibili vengano utilizzate per l'effettiva costruzione del modello [18]. I criteri di selezione delle variabili sono spesso basati sull'uso di test statistici univariati (e.g. Fisher e Wilcoxon) per valutare l'informazione delle feature e multivariati per misurare la correlazione esistente tra le variabili ed eliminare dall'insieme di analisi le variabili ridondanti [14].

Adizionalmente vengono usati, all'occorrenza, metodi per la gestione di problematiche specifiche, ad esempio lo SMOTE (Synthetic Minority Oversampling Technique) che permette di risolvere gli sbilanciamenti presenti tra i dati ottenuti dalle immagini [18], e tecniche di convalida incrociata per la validazione dei risultati [14].

2.6 Intelligenza artificiale in radiomica

2.6.1 Aspetti generali

Dopo aver evidenziato le difficoltà che si incontrano nell'elaborazione dei dati radiomici per mezzo di metodi statistici tradizionali, risulta chiara la necessità ed utilità nell'utilizzare un approccio basato sull'intelligenza artificiale per eseguire le medesime funzioni.

L'IA, infatti, rende molto più efficiente l'analisi di grandi moli di dati e il riconoscimento di *pattern* all'interno di essi. Tra le tecniche di intelligenza artificiale, quelle che in tempi recenti hanno subito un maggiore sviluppo e, quindi, ampio utilizzo sono basate sul *machine learning* (ML da qui in avanti), o apprendimento automatico, e nello specifico il *deep learning* (DL).

L'implementazione ed integrazione del ML nel workflow radiomico prevede i seguenti passaggi [19]:

- **Predisposizione dei requisiti.** Al contrario degli approcci classici, nel ML più grande è il data-set e migliori sono i risultati; al fine di aumentare il numero di dati sono eseguite tecniche di *data augmentation*, specialmente nell'ambito del deep learning. Questi stessi dati devono poi subire normalizzazioni, dato che il processo di estrazione spesso produce classi di dati prive di una scala di riferimento.
- **Sviluppo del modello.** I modelli di ML usati sono detti classificatori, ovvero algoritmi che suddividono i dati ricevuti in input in classi identificate da un'etichetta; di fatto sono una mappa dallo spazio delle features ad un insieme di etichette. I più diffusi in questo campo, appartenenti al tipo *supervised*, sono: *k-nearest neighbour*, regressioni logistiche, *naive Bayes* (bayesiani ingenui), *random forest*, *support vector machine*, *decision tree* e *neural networks* (reti neurali, NN da qui in avanti).
- **Validazione del modello.** Per entrare nella pratica clinica, l'analisi radiomica deve vedere i risultati ottenuti da studi pilota in piccola scala, che al momento caratterizzano la letteratura disponibile, validati da studi su larga scala e soprattutto tra istituti diversi (si vedrà in seguito il problema della validazione inter-istituto). Anche per questo scopo esistono diverse tecniche dette di *cross-validation* (*nested*, *leave-one-out*, *hold out*, *bootstrap*, *k-fold* e *random subsampling*).

- Valutazione della performance.** Solitamente la valutazione della performance del modello avviene attraverso il calcolo dell'*Area Under the Curve* (AUC) della curva *Receiver Operating Characteristic* (ROC). La curva caratteristica del ricevitore è un grafico che rappresenta la relazione tra falsi positivi e veri positivi di un classificatore binario e consente, quindi, di predirne la quantità dei primi sulla base del numero di test effettuati [3]. Tale metodo è, però, inefficace in presenza di sbilanciamenti nelle classi di dati con cui si lavora e rende necessario applicare altri controlli statistici quali il test t-student e quello signed-ranked di Wilcoxon [19]. Assieme alla ROC, vengono definite altre due misure della qualità di un test diagnostico ad essa collegate, le quali vengono spesso utilizzate in combinazione con la prima: specificità e sensibilità. La specificità, o *specificity*, descrive la capacità del test di identificare correttamente i soggetti sani e mantenere bassa la percentuale di falsi positivi, mentre la sensibilità, o *sensitivity*, svolge la funzione opposta, quantificando la capacità del test di identificare i veri ammalati e tenere bassa la percentuale di falsi negativi.

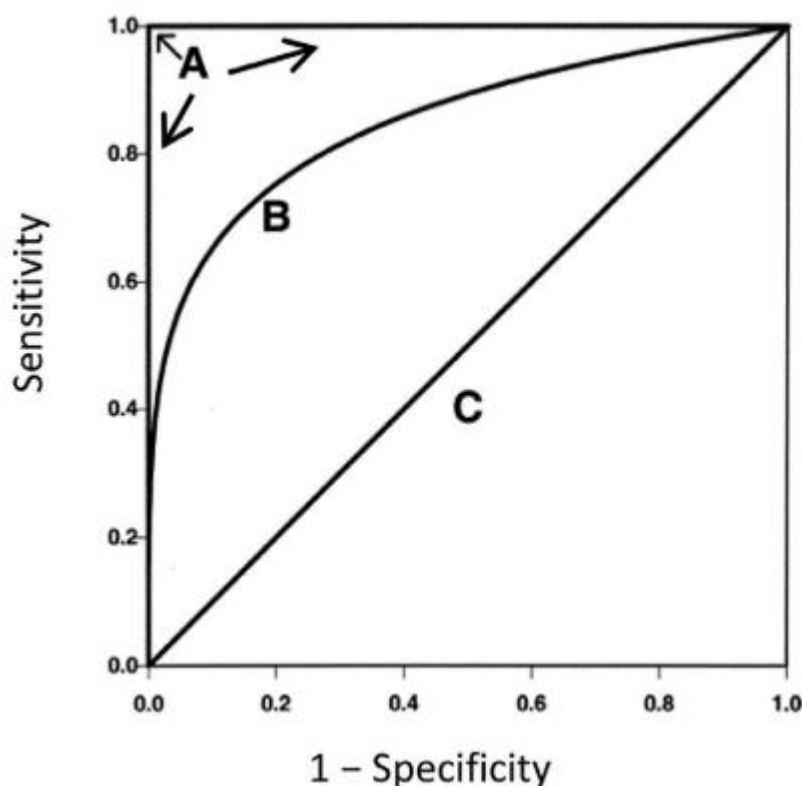


Figura 6: Esempi di grafico di ROC, si noti la relazione degli assi con sensibilità e specificità. Le curve A e C rappresentano i casi limite di un test ideale (solo veri positivi e veri negativi) e uno casuale (tanti veri quanti falsi positivi).

I difetti principali del machine learning che si riscontrano nell'applicazione alle immagini di diagnostica riguardano l'interpretazione del processo di apprendimento automatico ed i suoi risultati ed il rischio di overfitting del modello. Il primo problema consta nella difficoltà di

comprendere il processo “intelligente” che sta dietro all’analisi effettuata con il ML e nel conseguente confronto con le conclusioni tratte parallelamente dal personale medico; in sostanza risulta quasi impossibile per un radiologo “discutere” e trarre beneficio dal confronto con i risultati dell’analisi automatica. Nonostante ciò, il trade-off tra performance e interpretabilità è ancora vantaggioso verso il primo termine ed inoltre sono in corso di sviluppo modelli interpretabili basati sul deep learning noti come *explainable AIs*.

Il secondo problema invece è dovuto a caratteristiche intrinseche di questo tipo di approccio, il quale dimostra scarsa solidità nelle performance con dati non appartenenti all’insieme di training, circostanza frequente nel caso di malattie rare [20].

2.6.2 Reti Neurali e Deep Learning

L’uso delle reti neurali e del deep learning consente di ovviare a diverse problematiche che insorgono con approcci di ML più semplici, come già citato sopra nel caso delle explainable AIs. Nel campo della *Computer Vision*, ovvero quella branca dell’intelligenza artificiale che si occupa di sviluppare modelli in grado di simulare le prestazioni della visione umana, un modello di rete neurale con il quale vengono consistentemente ottenuti ottimi risultati e il cui uso è ormai associato è la *Convolutional Neural Network* (CNN, rete neurale convoluzionale). Il primo step del flusso radiomico in cui le CNNs trovano immediata integrazione è la segmentazione, dove in tempi recenti hanno consentito ad ottenere risultati promettenti [21]. Con questo tipo di reti si sono anche raggiunti i migliori risultati di interpretabilità attraverso il *Gradient-weighted Class Activation Mapping*, che consente di osservare il grado di attivazione e quindi rilevanza di specifici dati all’interno del set [20].

Il deep learning ha, inoltre, la potenzialità di eseguire in autonomia l’intero processo di analisi, viene perciò detto *unsupervised*, svolgendo l’estrazione e selezione delle features automaticamente e ripetendo il passaggio più volte con esiti diversi, integrandolo a tutti gli effetti nel training [22]. In questo senso si comprende come l’intelligenza artificiale sia uno strumento che può essere inserito all’interno del workflow radiomico, più precisamente nella fase finale di *pattern-recognition* e *data mining*, ma può anche costituire un mezzo a sé stante con cui eseguire congiuntamente tutte le fasi dell’analisi.

Ora che è stato illustrato interamente il processo che costituisce l’analisi radiomica si procede, nel capitolo che segue, con la presentazione di alcuni casi studio tratti dalla recente letteratura medica e radiologica, con cui è possibile apprezzare l’effettiva applicazione pratica della tecnologia ed i risultati che consente di ottenere.

3. APPLICAZIONI DI RADIOMICA IN ONCOLOGIA

In questo capitolo vengono presentati dei casi studio di applicazione di radiomica nella diagnosi e nella prognosi delle due patologie tumorali più diffuse nel mondo: il cancro al seno e il cancro del polmone. Gli articoli citati, provenienti da giornali e riviste scientifiche di medicina e radiologia, riportano recenti risultati ottenuti grazie all'analisi radiomica nello svolgere precisi compiti diagnostici e prognostici. Per ogni articolo si discutono i materiali e metodi utilizzati insieme ai risultati numerici e quantitativi ottenuti, al fine di poter trarre delle conclusioni riguardo la qualità degli esiti e la conseguente performance della radiomica, ma anche riguardo alcuni aspetti più tecnici della sua implementazione cruciali nel campo della ricerca e sviluppo.

3.1 Diagnosi e trattamento del cancro al seno

3.1.1 Rilevanza della patologia e metodi di diagnosi

Il cancro al seno rappresenta la quinta causa di morte nella popolazione mondiale, con circa 2,3 milioni di casi nel 2020 ed è, nello stesso anno, diventato il cancro più diagnosticato, superando il cancro ai polmoni. Attualmente è il tumore più frequente tra le donne, con una prevalenza del 24,5% ed è la principale causa di mortalità (15,5%). Almeno una donna su otto riceverà una diagnosi di cancro al seno nel corso della sua vita [23]. Considerando questi dati è facile dedurre l'importanza dello screening per questa patologia. Attualmente il metodo diagnostico principalmente usato è la mammografia digitale a campo pieno (*Full-field digital mammography*, FFDM), dalla quale è possibile descrivere forma, densità, contorno e presenza di calcificazioni che insieme ad altri segni clinici del paziente consentono di formulare la diagnosi tramite l'applicazione del BI-RADS (*Breast Imaging-Reporting and Data System*), un sistema di classificazione del tumore al seno riconosciuto internazionalmente. Il difetto principale della mammografia digitale è la scarsa sensibilità nel distinguere il tessuto mammario benigno da quello maligno quando è molto denso, fattore che aumenta il numero di falsi positivi [24]. La tecnica di imaging che viene solitamente usata per sopperire a questa mancanza è l'ecografia ad ultrasuoni (US), che consente anche di caratterizzare con più dettaglio le lesioni rilevate, riuscendo a identificare in media dai 2 ai 7 tumori in più ogni 1000 casi analizzati [23]. In generale, l'obiettivo dell'applicazione della radiomica in questo campo è quello di permettere di caratterizzare con più precisione le lesioni estraendo informazioni quantitative aggiuntive e mantenere in un giusto equilibrio il rapporto tra il numero di falsi positivi e quello di falsi negativi nelle diagnosi.

3.1.2 Radiomica per la previsione del sottotipo cellulare

La definizione del sottotipo nel cancro al seno ha importanti implicazioni terapeutiche sulla gestione della malattia. I principali sottotipi molecolari di tumore al seno sono: luminale A, luminale B, *human epidermal growth factor receptor 2 (HER2) – enriched* e triplo negativo. I tumori luminali rappresentano la maggioranza (70%) dei tumori al seno invasivi e rispondono bene alla terapia endocrina. I tumori HER2-enriched sono i candidati migliori per terapia con anticorpi. I tumori tripli negativi possono essere più aggressivi e difficili da trattare, ma alcuni di essi possono rispondere bene alla chemioterapia. Generalmente, i sottotipi di cancro al seno possono essere classificati solo dopo una biopsia, mediante test di *array* genetici o utilizzando marcatori immunostochimici (IHC).

In uno studio del 2019 un gruppo di ricercatori del Tianjin Medical University Cancer Institute and Hospital (Tianjin, Cina) ha studiato l'associazione tra le feature radiomiche estratte da immagini di mammografia digitale (DM) e i sottotipi cellulari di cancro al seno [18].

Lo studio svolge un'analisi retrospettiva di 331 donne a cui è stato diagnosticato un cancro al seno invasivo nel 2015, ognuna delle quali è rappresentata da una coppia di immagini acquisite con angolazioni diverse: con vista craniocaudale (CC) e con vista mediolaterale obliqua (MLO). La coorte include 29 soggetti tripli negativi, 45 HER2-enriched, 36 luminali A e 221 luminali B.

La segmentazione delle immagini è stata effettuata manualmente da personale specializzato, mentre programmi informatici automatizzati hanno calcolato 39 caratteristiche radiomiche:

- 1) caratteristiche morfologiche, come forma, dimensione, area, perimetro, descrittori di rotondità, concavità e coefficiente di Fourier;
- 2) caratteristiche statistiche ottenute dall'istogramma (primo ordine), come media, deviazione standard, asimmetria e curtosi
- 3) textures del secondo ordine per quantificare l'eterogeneità calcolata utilizzando la GLCM, come ad esempio contrasto, correlazione, energia, entropia, omogeneità, inerzia e momenti inversi.

Di queste 39 features, rotondità, concavità, media di grigio e correlazione sono state individuate come più rilevanti.

Il test prevede tre classificazioni binarie dei sottotipi: triplo negativo vs. non triplo negativo, HER2-enriched vs. non arricchito e luminale (A+B) vs. non luminale.

Per la classificazione è stato usato uno schema di ML di tipo naive Bayes. Il LASSO è stato utilizzato per selezionare le caratteristiche più predittive per i classificatori, infine, le prestazioni di classificazione sono state valutate in base all'AUC e la sensibilità.

I risultati di valutazione sono riportati nella tabella seguente, distinti rispetto al tipo di acquisizione usata per il test:

Data	Triple-negative Vs. Non- triple-negative	HER2-enriched Vs. Non- HER2-enriched	Luminal Vs. Non-Luminal
CC view	0.695 / 0.634	0.741 / 0.725	0.755 / 0.785
MLO view	0.853 / 0.772	0.756 / 0.702	0.791 / 0.782
CC and MLO	0.865 / 0.796	0.784 / 0.748	0.752 / 0.788

Tabella 1: Valori di AUC / sensitivity dei tre test eseguiti sulle immagini divise per angolazione e totali (da [18]).

Consultando la tabella si può concludere che dalla mammografia è possibile estrarre informazioni riguardanti il sottotipo cellulare della massa analizzata con ottimi livelli di accuratezza, specialmente combinando angoli di acquisizione diversi. L'esito ottenuto costituisce una prova sperimentale dell'ipotesi radiomica esposta nell'introduzione.

3.1.3 Diagnosi differenziale di malignità da mammografia ed ecografia

In questa sezione si comparano due studi in cui si esegue una diagnosi differenziale di tumore benigno/maligno al seno a partire da immagini di mammografia nel primo ed ecografia nel secondo.

Nel primo studio, risalente al 2018, Zhiming et al. sfruttano l'analisi di textures per l'estrazione di informazioni utili come supporto nella differenziazione di lesioni benigne e maligne a partire da immagini di mammografia digitale [24].

Anche questo è uno studio retrospettivo su una coorte di 302 donne diagnosticate con tumore al seno, di cui 226 con tumore maligno e 76 con lesioni benigne. Tutte le pazienti sono state diagnosticate con il metodo BI-RADS (categoria $\geq 4B$ per il caso maligno e $< 4A$ per quello benigno) di tumore della mammella unilaterale, non derivante da metastasi. La diagnosi è stata eseguita con il supporto delle immagini di mammografia e verificata poi con un esame istologico. Al fine di valutare l'utilità di integrare l'analisi delle texture nella diagnosi basata sull'imaging, i ricercatori hanno comparato le prestazioni di un modello radiomico con quelle ottenute da personale specializzato usando BI-RADS e con quelle ottenute applicando sinergicamente l'osservazione delle immagini di DM e la texture analysis (3 test totali).

Entrando nei particolari del processo di implementazione, la regione di interesse (ROI) è stata delineata manualmente. Si è proceduto poi con il calcolo automatico delle features per mezzo del software Omni-Kinetics. Le textures estratte sono del primo ordine (istogramma) e del secondo (GLCM e GLRLM). I dati sono stati trattati con una regressione logistica binaria ed il risultato è stato valutato calcolando l'area sotto la ROC (AUROC).

Le differenze più significative delle texture tra tumori benigni e maligni rilevate riguardano: 19 features dell'istogramma, 5 nella GLCM e 11 nella GLRLM.

Le AUROCs di diagnosi basata sul BI-RADs, analisi delle texture e sulla combinazione dei due metodi sono rispettivamente di 0,873, 0,863 e 0,961 (CI al 95% per tutte). È evidente che, il caso combinato presenta un'AUROC maggiore di quelle ottenute usando le tecniche da sole, le quali AUROC non hanno differenze significative.

Nel secondo studio, pubblicato nel 2023 e svolto dalla Breast Imaging Division dell'IEO (Istituto Europeo di Oncologia, Milano, Italia), Nicosia et al. valutano le prestazioni dell'analisi radiomica e dell'uso di un CAD per lo stesso scopo del paper precedente, ma questa volta su immagini di ecografia ad ultrasuoni (US) [23].

Qui la base di partenza è un database di 365 immagini ad US di lesioni mammarie provenienti da 362 pazienti che sono state sottoposte a biopsie per lesioni sospette (BI-RADS > 3) tra gennaio 2020 e marzo 2022. Le immagini sono state, in primo luogo, revisionate da quattro radiologi esperti, i quali le hanno classificate nuovamente secondo il sistema BI-RADS, considerando la soglia di malignità al valore 4A (escluso, uguale al precedente studio).

La segmentazione è stata eseguita manualmente da tre radiologi differenti con il programma LifeX.

Il CAD S-Detect, basato su un algoritmo di deep learning per valutare le lesioni mammarie ecografiche, è stato utilizzato per ottenere una suddivisione delle lesioni in “possibilmente benigne” e “possibilmente maligne”. Per ogni lesione, il radiologo, dopo l'acquisizione di immagini in B-mode sia longitudinale che trasversale, ha posto un segno nel centro della lesione: attorno ad essa, lungo il bordo della massa, il sistema fornisce automaticamente una regione di interesse (ROI) e assegna un punteggio di malignità (in termini di probabilità). Il risultato della biopsia istopatologica è stato utilizzato come *gold standard* per definire l'accuratezza diagnostica, analogamente all'altro studio.

Ancora una volta per selezionare le features radiomiche è stato implementato il LASSO, combinato con una *10-fold cross validation*.

Successivamente le immagini sono state assegnate in modo casuale ad un set di addestramento e ad uno di test del modello con un rapporto di 70 a 30 (255 immagini per l'addestramento e 110 per il test). Da questi due insiemi, dei modelli di regressione logistica univariabile e multivariabile per la previsione della malignità sono stati costruiti eseguendo il fitting (adattamento) con il primo e la valutazione col secondo.

La performance del modello costruito è stata, infine, valutata con l'usuale calcolo dell'AUROC, della sensibilità e della specificità.

I risultati finali sono:

Model (Training Group)	AUC	SE	SP
Adjusted Radiomic + S-Detect	0.887 (0.840–0.933)	0.854 (0.771–0.916)	0.802 (0.716–0.873)
Adjusted Radiomic + BI-RADS best	0.883 (0.839–0.927)	0.854 (0.854–0.771)	0.764 (0.672–0.841)
Radiomic + S-Detect + BI-RADS best	0.914 (0.876–0.951)	0.854 (0.771–0.916)	0.849 (0.766–0.911)

Tabella 2: Valori di AUC, sensibilità (SE) e specificità (SP) dei tre modelli testati. Tra parentesi è riportato l'intervallo di confidenza al 95%.

Gli studi presentati ben supportano l'applicazione della radiomica allo scopo prefissato di eseguire diagnosi di malignità del tumore al seno. In particolare, entrambi dimostrano che la combinazione di un modello clinico-radiomico raggiunge le prestazioni migliori, mentre il secondo consente anche di osservare che l'uso di un CAD automatico aumenta ulteriormente la qualità dei risultati, permettendo di contrastare la variabilità inter-operatore.

3.2 Diagnosi e prognosi del cancro ai polmoni

3.2.1 Aspetti clinici

Il polmone rappresenta la seconda sede di cancro più frequente (dal 2020, in precedenza era al primo posto) e la prima per numero di morti ogni anno. Circa l'85% dei tumori del polmone ricade nella classe dei carcinomi polmonari non a piccole cellule (Non-Small Cell Lung Cancer, NSCLC), che a sua volta presenta le seguenti istologie con maggiore frequenza: adenocarcinoma, carcinoma squamocellulare e carcinoma a grandi cellule. Nella medicina tradizionale il metodo di stadiazione del cancro polmonare più comune (usato anche per altri tumori) è il *Tumor-Nodes-Metastasis* (TNM), il quale è basato su tre criteri anatomici: la dimensione della massa tumorale, il coinvolgimento dei linfonodi regionali (nello stesso distretto corporeo della massa) e la presenza di metastasi lontane. Per gli stadi iniziali di cancro al polmone, quindi, l'escissione completa dei linfonodi sospetti, seguita da una valutazione microscopica, è il metodo più accurato per determinare la presenza di metastasi. Con questa tecnica le prestazioni diagnostiche migliori si hanno analizzando un numero elevato di linfonodi ed apportando, di conseguenza, maggiori traumi per i pazienti [25]. Considerando ciò, risulta evidente la necessità di sviluppare metodi di indagine e valutazione dei linfonodi meno invasivi di quelli usati oggi.

Parallelamente, i recenti progressi nella tecnologia di sequenziamento del DNA e dell'RNA hanno portato ad una prima comprensione di quali cambiamenti genomici determinano il fenotipo del cancro al polmone. Questi strumenti genomici sono ora disponibili a costi inferiori al passato, aprendo ulteriori possibilità per terapie personalizzate per il paziente e portando a

cambiamenti significativi nella pratica clinica, che è ora guidata dalla ricerca dei biomarcatori, come la mutazione EGFR e la traslocazione ALK per il NSCLC.

Al momento della diagnosi, quasi tutti i pazienti sono sottoposti a tecniche di imaging, come la tomografia computerizzata (CT) e la tomografia a emissione di positroni (PET), per la stadiazione e per escludere la presenza di lesioni secondarie. L'associazione tra i descrittori ottenuti dalle immagini e talune caratteristiche della patologia è ormai consolidata, ad esempio la presentazione radiologica degli adenocarcinomi polmonari include quasi sempre un ampio spettro di aspetti comuni sia nei noduli che nelle masse solide. Inoltre, molti descrittori ricavati dalle immagini permettono di riconoscere i diversi tipi istologici, i principali sono densità, forma, margini, la *ground-glass opacity* (GGO), la presenza di cavitazioni e necrosi ed il broncogramma aereo.

Tali considerazioni aprono la strada alla possibilità di progresso nella direzione di integrazione delle informazioni genomiche e fenotipiche fornite dalle nuove tecnologie di sequenziamento con le immagini di diagnostica, le quali promettono grande utilità nel processo di stratificazione [26].

3.2.2 Assegnazione di un radiomic-score per valutare il grado metastatico del NSCLC

In uno studio risalente al 2020 dell'Istituto Oncologico Europeo (IEO) si è ricercata un'associazione tra le caratteristiche quantitative di immagini da CT e due fattori prognostici di fondamentale importanza del NSCLC: lo status dei linfonodi del paziente e il tasso di sopravvivenza (*Overall Survival*, OS) [25].

La coorte consiste in 270 pazienti sottoposti a linfadenectomia, ovvero un'operazione di drenaggio del linfonodo, ed è stata suddivisa in due parti, una da 180 pazienti per il training e l'altra da 90 pazienti per la validazione del modello. In partenza sono state estratte 881 features radiomiche dal singolo individuo, il modello è stato poi costruito su tre selezioni di textures diverse, ottenute dall'insieme iniziale: la prima con l'algoritmo di Filtered Back Projection (FBP), la seconda con l'algoritmo di Iterative Reconstruction (IR) e l'ultima con entrambi gli algoritmi combinati.

I due test prevedevano la comparazione delle performance di un modello clinico, ovvero basato su evidenze cliniche, con un modello interamente radiomico ed un terzo modello misto clinico-radiomico al fine di valutare le potenzialità dell'analisi radiomica.

Nel primo test, ovvero quello in cui si indagava la positività dei linfonodi alla metastasi, le features di maggiore rilevanza per il training sono state il ClusterShade della GLCM, il 70^{imo} percentile dei valori di intensità dell'istogramma ed il massimo diametro volumetrico della

lesione. A partire da queste si è definito il radiomic-score usato per la determinazione della positività del linfonodo. Il modello clinico invece, ha considerato come fattori determinanti per la positività la variabilità del sito di lesione e la dimensione del nodulo. Nell'immagine che segue sono riportate le ROC (e relative AUC), dei tre modelli predittivi costruiti, ottenute rispettivamente con il training set ed il validation set:

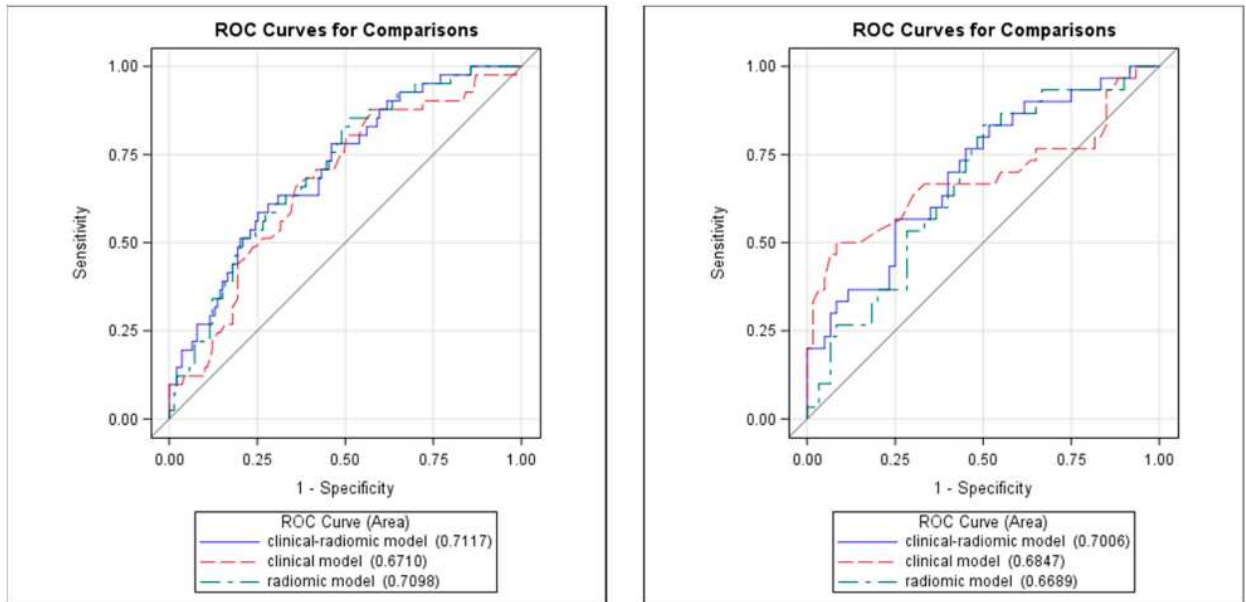


Figura 7: curve ROC per la previsione di positività dei linfonodi del training set (sinistra) e validation set (destra) (da [25]).

Dalle immagini si può dedurre che il modello clinico-radiomico combinato ha performance equiparabili a quelle dei due modelli singoli. Nello studio viene imputata una preselezione troppo stringente delle features per spiegare lo scarso miglioramento ottenuto.

Considerando gli algoritmi singolarmente, invece, risulta che il modello basato su IR con un AUC di 0.76 è più efficace di quello FBP con AUC di 0.61.

Nel secondo test si è stimato l'OS usando un modello radiomico allenato su 6 features selezionate con un'analisi LASSO dei set ottenuti da IR e FBP ed un modello clinico che andava a valutare i seguenti parametri medici: il lato ed il sito del tumore e l'indice di classificazione patologica secondo il TNM (indicato come pT). Nella tabella seguente sono riportate le prestazioni di tutti i modelli testati:

Model	Algorithm	C-Index (95%CI) Training Set	C-Index (95%CI) Validation Set
Radiomic	FBP + IR	0.73 (0.66–0.80)	0.59 (0.47–0.70)
Clinical (side, site, pT)	FBP + IR	0.78 (0.71–0.85)	0.53 (0.41–0.65)
Clinical–radiomic	FBP + IR	0.77 (0.69–0.84)	0.57 (0.46–0.69)
Radiomic	FBP	0.69 (0.61–0.77)	0.61 (0.46–0.76)
Clinical (site, pT)	FBP	0.66 (0.57–0.75)	0.57 (0.43–0.72)
Clinical–radiomic	FBP	0.70 (0.63–0.78)	0.61 (0.46–0.75)
Radiomic	IR	0.82 (0.66–0.98)	0.82 (0.53–1.00)
Clinical (site)	IR	0.67 (0.50–0.85)	0.65 (0.38–0.91)
Clinical–radiomic	IR	0.84 (0.68–1.00)	0.89 (0.60–1.00)

Tabella 3: Indice di concordanza (C-Index) dei vari modelli costruiti per i rispettivi set, distinti anche in base all’algoritmo di selezione. Si può notare che sono stati testati diversi modelli clinici con sottoinsiemi dei fattori considerati (da [25]).

Dai risultati è evidente che la migliore performance è stata ottenuta con il modello combinato clinico-radiomico basato sulla selezione con algoritmo IR e che, complessivamente, i modelli radiomici hanno sovraperformato quelli clinici. Ciò permette di concludere che l’uso di un radiomic-score può acconsentire a migliori decisioni prognostiche nel caso in cui il modello predittivo sia stato correttamente implementato e che le performance di tale modello dipendono considerevolmente dagli algoritmi utilizzati per costituire il training set.

4. LIMITI E PROSPETTIVE DELLA RADIOMICA

4.1 I limiti dell'analisi radiomica

4.1.1 Situazione attuale della tecnologia

L'articolo "*Radiomics extracting more information from medical images using advanced feature analysis*" di Lambin et al. (2012) è stato il primo paper pubblicato a parlare di radiomica e "*Radiomics: Images Are More than Pictures, They Are Data*" di Gillies et al. (2016) è il paper più citato del settore, con 3024 citazioni ad agosto 2022 [27]. Il termine *Radiomics*, ed il concetto che si porta dietro, ha, quindi, poco più di dieci anni. Sarebbe dunque improprio definire la radiomica come un campo della ricerca medica maturo e vicino all'inserimento nella pratica clinica su scala globale. È più proprio parlare di un ambito di ricerca avanguardistico ancora in fase di sperimentazione, che ha visto in tempi recenti un grosso slancio legato soprattutto ai fattori di seguito elencati:

- **Innovazioni dei dispositivi medici.** L'hardware ha subito importanti miglioramenti aumentando la qualità delle immagini e acconsentendo all'uso di acquisizioni combinate quali CT/PET, MRI/PET.
- **Innovazioni degli agenti di imaging.** L'individuazione di biomarcatori, radio-traccianti e molecole *imaging-probe* (sonde) permette di descrivere con accuratezza precisi stati biologici, fattore fondamentale per lo sviluppo di metodi radiomici.
- **Innovazioni software nell'imaging.** L'integrazione di CAD ha permesso di intraprendere un percorso basato sull'analisi quantitativa di immagini, che, accompagnata dallo sviluppo dell'AI apre la strada al data-mining in medicina [10].

Il trend di crescita appena esposto può essere apprezzato dal grafico in figura 8, dove viene mostrato il numero di papers contenenti la parola chiave *radiomics* pubblicati ogni anno.

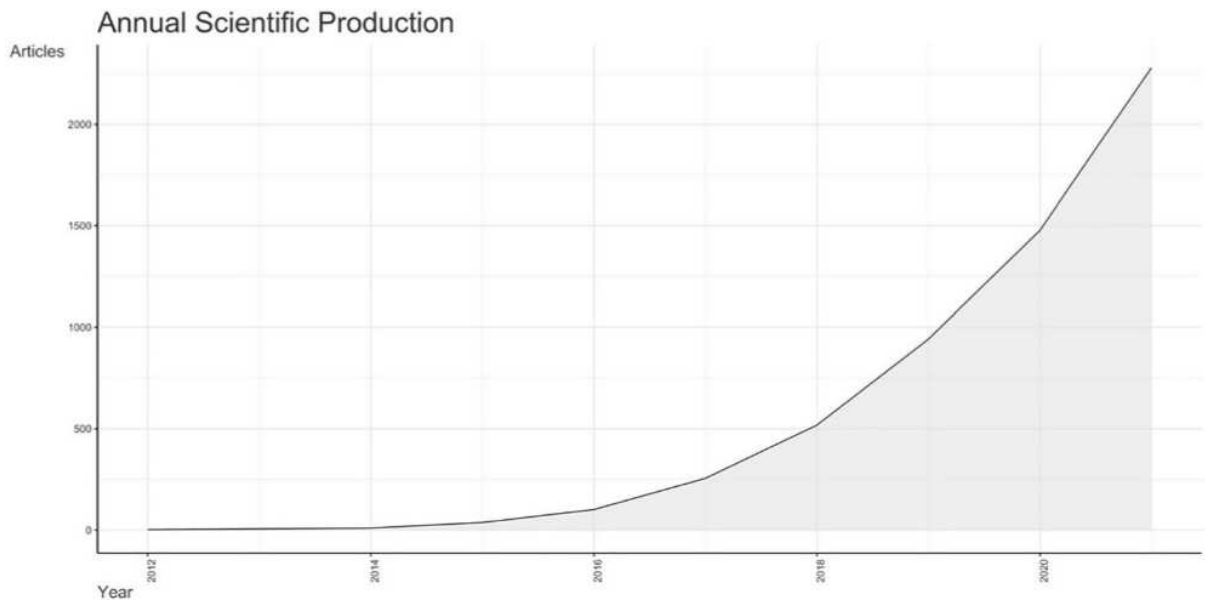


Figura 8: Papers con keyword “radiomics” nel tempo, dal 2012 ad oggi (2023) (da [27]).

Ciò detto, il settore della radiomica presenta alcuni limiti piuttosto importanti che ne compromettono fortemente lo sviluppo, i quali possono essere categorizzati in tre classi principali:

- variabilità dei risultati ottenuti dai modelli rispetto a condizioni interne alla loro costruzione
- mancanza di standard internazionali che mina la ripetibilità dei test e la concordanza inter-istituto
- assenza di databases strutturati che consentano un’implementazione su larga scala del data mining sfruttando la potenza dell’AI.

4.1.2 Variabilità dei risultati

Come già riscontrato negli studi citati precedentemente, i modelli predittivi radiomici presentano numerosi *bias*, che si traducono in una forte variabilità della performance, dovuti a fattori interni all’analisi come il processo di selezione delle textures, l’esito della segmentazione (vedi figura) ed altre variabili relative all’acquisizione dell’immagine.

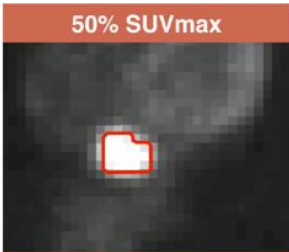
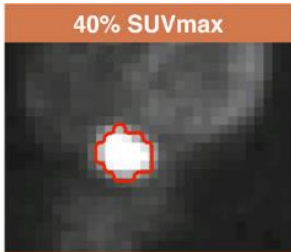
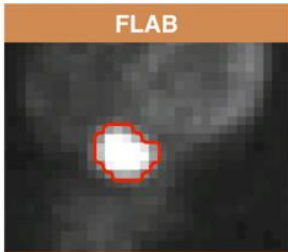
	50% SUVmax	40% SUVmax	FLAB		
					
PET feature	Value	Value	Change	Value	Change
Volume	26.59	34.61	+30%	40.36	+51%
Entropy	3.64	3.89	+7%	3.98	+10%
Total energy	$4.72 \cdot 10^4$	$5.28 \cdot 10^4$	+12%	$5.57 \cdot 10^4$	+18%
Homogeneity	0.29	0.25	-12%	0.24	-17%
Dissimilarity	2.97	3.60	+21%	3.95	+33%
Zone percentage	0.35	0.39	+9%	0.39	+10%
Size-zone variability	0.36	0.38	+6%	0.38	+4%
Intensity variability	0.09	0.08	-15%	0.07	-18%

Figura 9: Sono riportate le differenze di valore di alcune features in funzione della segmentazione effettuata su immagini PET (da [14]).

In uno studio condotto dall'Unità di Fisica Medica dell'Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST, Mendola, Italia) viene utilizzata l'analisi della texture per valutare come l'angolo di acquisizione della tomosintesi digitale del seno (*Digital Breast Tomosynthesis*, DBT) possa influenzare l'acquisizione di immagini del parenchima mammario [28]. I sistemi DBT utilizzano una serie sofisticata di parametri di acquisizione per generare un set di immagini, tra questi l'angolo di acquisizione è uno dei più importanti. In questo caso l'analisi delle textures viene usata in modo sostanzialmente inverso a quello tradizionale, con lo scopo di studiare appositamente come variazioni dell'input si riflettano sull'output prodotto. Lo studio è stato condotto su un gruppo di 34 pazienti (per un totale di 36 immagini). Le features estratte sono state analizzate con il software MaZda variando, in prove diverse, il numero di bit di codifica dei livelli di grigio.

First-Order Features (Histogram)	Mean
	Variance
	Skewness
	Kurtosis
<hr/>	
Second-Order Features (GLCM)	Angular Second Moment
	Contrast
	Correlation
	Sum of Squares
	Inverse Difference Moment
	Sum Average
	Sum Variance
	Sum Entropy
	Entropy
	Difference Variance
Difference Entropy	

Tabella 4: Features estratte nello studio, distinte in base all'ordine (primo o secondo).

Per misurare le differenze di valore delle texture ottenute con i due diversi angoli è stato usato il test signed-rank di Wilcoxon, mentre per misurare la correlazione il rank test di Spearman, producendo i risultati seguenti, espressi come differenze tra valori medi:

		% Difference 40°-15° Acquisition Angle ($p < 0.05$)
<hr/>		
First-Order Features (Histogram)		
<hr/>		
	Contrast	+50%
	Correlation	-5%
Second-Order Features (GLCM)	Inverse Difference Moment	-14%
	Difference Variance	+45%
	Difference Entropy	+17%

Tabella 5: Differenze relative (esprese in percentuale) più rilevanti riscontrate nello studio (da [28]).

Dalla tabella si osserva che le features del primo ordine non hanno subito alcuna modifica rilevante (sono assenti). Diverse texture derivate dalla GLCM hanno, invece, subito importanti variazioni dimostrando la dipendenza dei valori delle texture calcolate da immagini DBT dall'angolo di acquisizione.

Un altro tipo di *inconsistency* che si rileva con l'analisi radiomica riguarda il diverso esito predittivo di modelli basati su piattaforme di feature extraction differenti. Nel 2022, ad esempio, l'IEO ha svolto uno studio comparativo di due software di estrazione delle texture, PyRadiomics (PyR) e SOPHiA Radiomics (SR), nel prevedere l'esito del trattamento radioterapico in pazienti con metastasi cerebrali del NSCLC [29].

Per il monitoraggio delle metastasi cerebrali (*Brain Metastasis*, BMs), la radiomica si basa principalmente sull'analisi di immagini di MRI, con cui sono stati ottenuti diversi risultati positivi nella previsione dell'origine delle BMs (da NSCLC e melanoma) [30], nella

differenziazione delle BMs da altre neoplasie (tumori gliali) [31] e nella valutazione della risposta al trattamento [32].

I ricercatori hanno deciso di utilizzare le impostazioni di base del software PyR, in modo da favorire la riproducibilità e il confronto, mentre con il programma SR, che non dispone di impostazioni di default, sono state calcolate textures di: GLCM, GLDZM, GLRLM, GLSZM, istogramma di intensità, intensità locale, morfologia, NGLDM, NGTDM, statistica e istogramma volume-intensità.

Al fine di valutare la variabilità prognostica dei modelli predittivi, si sono simulati due scenari diversi: nel primo ogni software ha usato il proprio set di textures ricreando il caso di ospedali diversi che effettuano analisi indipendenti; nel secondo entrambi i software hanno lavorato solo con il sottoinsieme delle textures condivise, in modo da confrontare precisamente le differenze di prestazione.

I modelli sono, quindi, stati usati per predire Local Control (LC, un indice di evoluzione del tumore primario), Distant Progression (DP, equivalente del LC per le metastasi) e Overall Survival (OS), producendo i seguenti risultati espressi come Concordance Index:

Test	Software	C-index set distinti	C-index set comune
Local Control	SR	0.70	0.69
	PyR	0.63	0.63
Distant Progression	SR	0.58	0.58
	PyR	0.55	0.57
Overall Survival	SR	0.64	0.61
	PyR	0.63	0.54

Tabella 6: Indice di concordanza dei modelli basati su SR e PyR nella previsione di LC, DP e OS ottenuti con il set di features proprie e quello di sole features in comune ([29]).

Complessivamente si può osservare che i modelli radiomici basati sul software SR sono sempre performanti nel calcolo del LC. Invece, le analisi della DP hanno rivelato che non esiste un modello nettamente migliore allo scopo. In conclusione, lo studio rivela che, oltre alla conferma della scarsa robustezza predittiva di modelli basati su diverse piattaforme, un numero maggiore di textures estratte non implica necessariamente migliori prestazioni. Infatti, nel test dove i due software hanno lavorato ognuno con il proprio set, SR ha performato meglio di PyR, nonostante il numero di textures usate da PyR fosse quasi sei volte quello di SR.

4.1.3 Standardizzazione e strutturazione dei dati

Il problema principale che impedisce alla radiomica di essere inserita a tutti gli effetti in un *Clinical Decision Support System* (CDSS), e quindi nella pratica clinica, è l'assenza di standard che si verifica a diversi livelli:

- **Standardizzazione dell'acquisizione di immagini e della gestione dei dati.** Nella sezione 4.1.1 è stato presentato un caso studio dove veniva evidenziata la dipendenza dell'analisi da un fattore interno all'acquisizione dell'immagine. In aggiunta, istituti diversi operanti con macchinari diversi produrranno a loro volta risultati sensibilmente incongruenti, rendendo necessaria una regolamentazione anche al livello di acquisizione e rappresentazione dei dati.
- **Standardizzazione del workflow radiomico**, specie nella feature extraction. Sempre nella sezione 4.1.1 e in alcuni studi riportati nel testo, si è visto come l'uso di software diversi o l'implementazione differente dello stesso software produca sensibili variazioni degli esiti ottenuti.
- **Standardizzazione nella rappresentazione dei risultati**, la quale rende possibile verificare, da parte di istituti diversi, i risultati di uno studio.

Tutti questi aspetti rendono proibitiva la ricerca inter-istituto e la ripetibilità dei test, rallentando lo sviluppo della tecnologia.

In tempi recenti, vi sono comunque stati ingenti sforzi da parte di diverse istituzioni nell'inizializzare il processo di standardizzazione. In campo oncologico il National Cancer Institute (NCI) ha creato un programma di cancer imaging con il quale ha istituito il *Quantitative Imaging Network* (QIN), la Radiological Society of North America (RSNA) insieme al National Institute for Biomedical Imaging and Bioengineering ha fondato la *Quantitative Imaging Biomarkers Alliance* (QIBA) [13], il Dipartimento di Immagini e Scienze Radiologiche dell'IEO ha creato un consiglio, il *Radiomic Board*, appositamente per la ricerca e sviluppo nel settore.

L'istituto di ricerca europeo BBMRI-ERIC (Biobanking and BioMolecular resources Research Infrastructure-European Research Infrastructure Consortium) ha sviluppato il *Minimum Information about BIobank data Sharing* (MIABIS), una risorsa per la gestione dei dati intra-istituto con cui sono definite le informazioni minime necessarie per iniziare una collaborazione tra banche di dati biologici, le *biobanks*, e permettere lo scambio di dati e campioni. In seguito, sono stati eseguiti studi per l'espansione ed il miglioramento di tale strumento [33].

Un'ultima iniziativa meritevole di menzione è l'*Imagin Biomarker Standardisation Initiative* (IBSI). Nata come collaborazione internazionale indipendente nel 2016, ha rilasciato nel 2020

il paper con cui ha sancito la conclusione del primo capitolo della sua missione: la standardizzazione delle features radiomiche comunemente utilizzate [34]. Tale lavoro ha stabilito valori di riferimento per 169 features, ha creato uno schema standard di elaborazione delle immagini radiomiche e ha sviluppato linee guida per il reporting per gli studi radiomici. È piuttosto evidente da questa lista di iniziative differenti quanto sia grande l'interesse, accompagnato da numerosi investimenti, degli istituti di ricerca e di sanità per questo settore, e ciò fa senz'altro presupporre che molti dei problemi saranno con ottima probabilità superati.

4.2 Prospettive future

4.2.1 Direzioni della ricerca

Il grafico in figura 8 (sezione 4.1.1) e le numerose iniziative citate nei paragrafi precedenti fanno intendere che il mondo della ricerca, ma anche quello della sanità in generale, nutre un forte interesse nei confronti di questa tecnologia e mostra una concreta volontà di renderla al più presto una parte effettiva della pratica clinica, specie del trattamento dei tumori.

In particolare, la ricerca sta ponendo maggiore attenzione su due obiettivi principali: l'applicazione del deep learning ([35]) e l'integrazione con la genomica.

La sfida principale che si pone davanti al primo obiettivo è la necessità di grandi moli di dati, soddisfabile solo con la costruzione di infrastrutture digitali che consentano l'immagazzinamento delle immagini ottenute da istituti, aziende ospedaliere e centri di ricerca separati per svolgere analisi su larga scala. Per fare ciò si devono superare problemi legati alla privacy e alla condivisione dei dati personali, oltre che alla costruzione di database e di protocolli di condivisione.

Il secondo obiettivo si traduce letteralmente nel termine composto Radiogenomica, la concreta fusione di genomica e radiomica nella comprensione e descrizione delle patologie oncologiche al fine di sviluppare terapie mirate su misura del paziente [26].

4.2.2 Integrazione con altri dati clinici

L'integrazione delle informazioni ottenute dall'analisi radiomica con quelle ottenuti da altri esami clinici o di rilevanza medica è alla base dello sviluppo della medicina personalizzata.

Di seguito si riportano degli esempi di tale integrazione per comprendere in quali modi e a che livelli possa avvenire.

In un'ulteriore espansione del MIABIS, Garau et al. propongono un modello di database che mira a standardizzare le raccolte di dati in cui sono coinvolti sia campioni biologici che digitali enfatizzando le relazioni causa-effetto tra i fattori poiché di grande rilevanza clinica. Per testare il modello in un contesto realistico, viene scelto il caso studio *Continuous Observation of SMOKing Subjects* (COSMOS), un programma di screening e follow-up decennale del cancro al polmone svolto su più di 5000 soggetti [36].

Partendo da estensioni preesistenti ([33]), lo studio effettua la progettazione concettuale di un database in cui query (interrogazioni) specifiche e non ambigue sono eseguibili e consentono la descrizione completa del repository COSMOS ed il conseguente uso per la ricerca nell'elaborazione delle immagini e nella generale gestione del cancro del polmone.

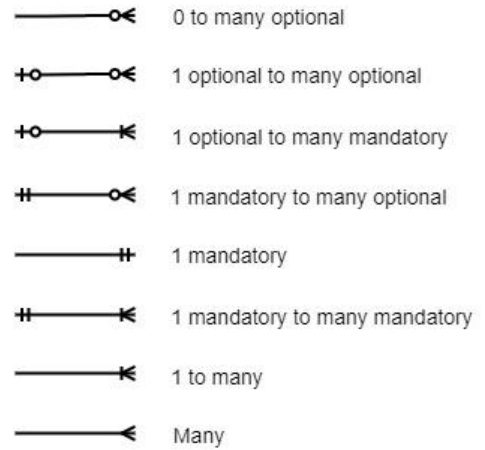


Figura 10: Legenda delle cardinalità delle relazioni nel diagramma in figura 13.

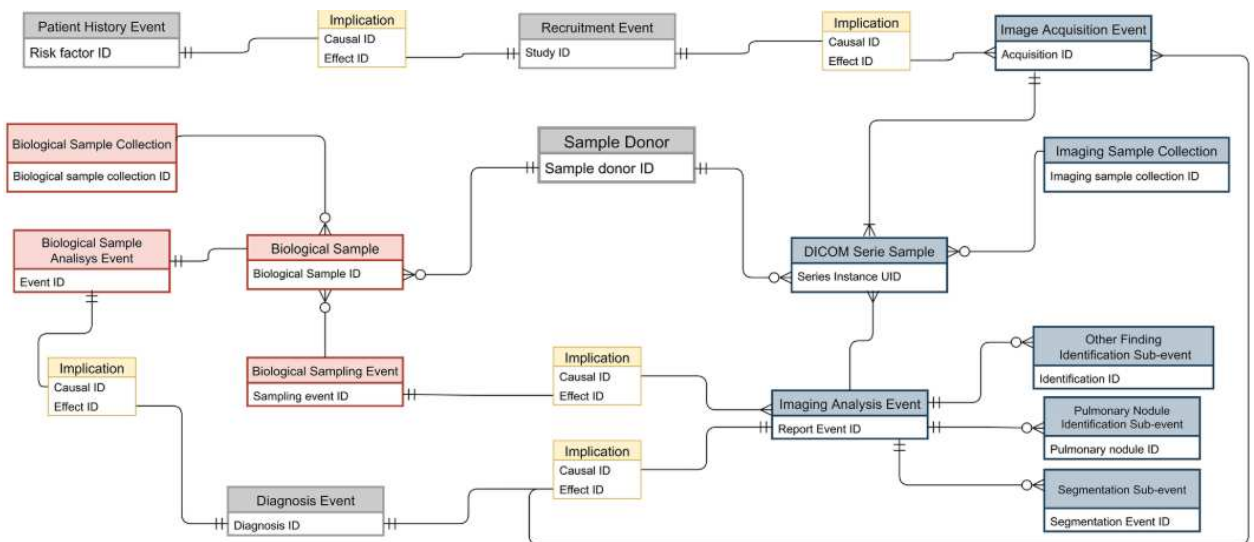


Figura 11: Diagramma relazionale del database COSMOS sviluppato in [36]. Sono riportate le tabelle di entità e relazioni con le rispettive primary key. Le entità sono distinte per colore in base alla loro natura, le relazioni (Implication) sono di tipo causa-effetto.

Nello studio europeo riportato in 3.1.3 ([23]), gli stessi ricercatori sviluppano un nomogramma che prende in considerazione: un *radiomic-score*, informazioni fornite dal CAD, il BI-RADS ed altri dati clinici. Un nomogramma è uno strumento di rappresentazione grafica di una formula matematica complessa che, in oncologia, mette in relazione dati clinici di diverso tipo e consente di effettuare diagnosi e prognosi, più accurate di quelle dei metodi tradizionali come il *Tumor-Node-Metastasis* (TNM) [37].

Il risultato è uno strumento di supporto prognostico del tumore capace di integrare dati di diverso tipo e facile da utilizzare dal personale (non necessita di formazione aggiuntiva).

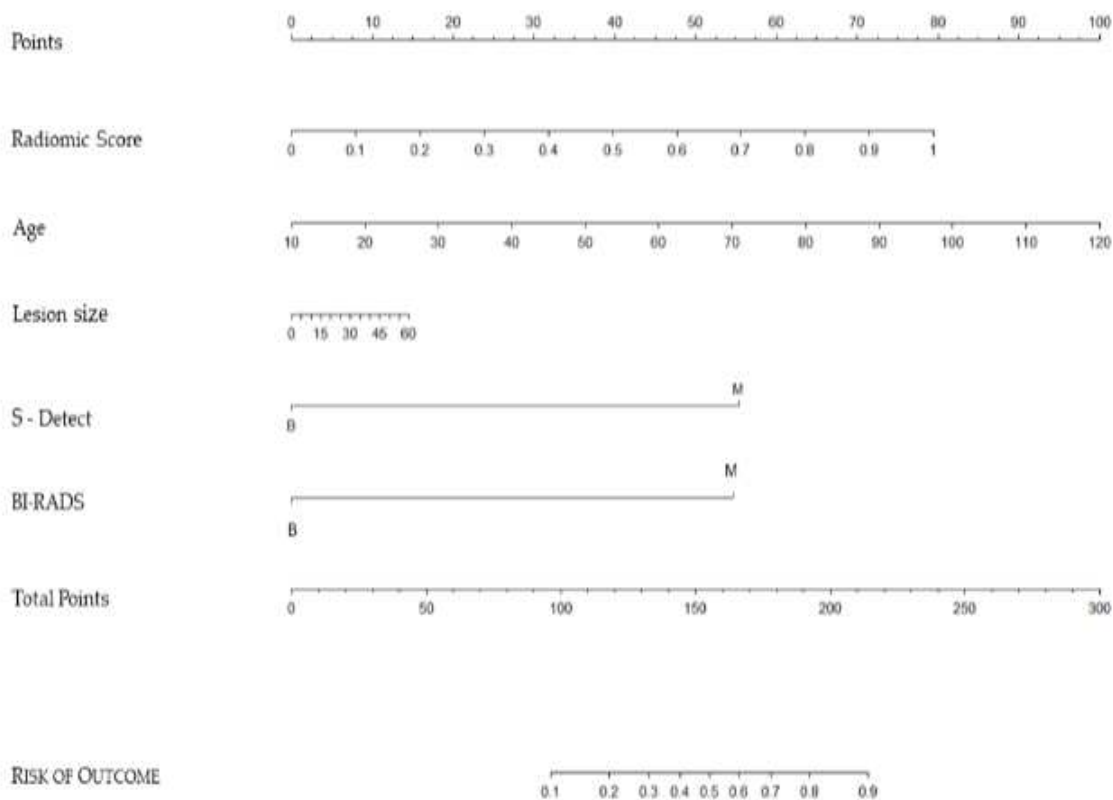


Figura 12: Nomogramma di [21]. La Lesion size è espressa in millimetri, S-Detect e BI-RADS forniscono un termine di possibile malignità di tipo probabilistico, in pratica un valore compreso tra gli estremi B (sicuramente benigno) e M (sicuramente maligno).

Questo nomogramma permette di assegnare un punteggio complessivo ad un paziente da confrontare con la scala “Total Points” per definire il grado di malignità del tumore. La costruzione del punteggio totale avviene sommando i punteggi dei singoli fattori. Il punteggio di un singolo fattore è definito attraverso la scala “Points” ed è ottenuto proiettando il valore del fattore (un punto nel corrispettivo segmento) lungo la suddetta scala. Una volta definito il punteggio totale è possibile predire il rischio effettivo di malignità (tra 0.1 e 0.9) proiettando il punteggio sulla scala nominata “RISK OF OUTCOME”.

Lo studio appena presentato fornisce, quindi, un esempio pratico di integrazione di risultati radiomici in uno strumento di medicina tradizionale ampiamente usato nella pratica clinica oncologica.

4.2.3 Uso del deep learning

Come si è discusso nella sezione 2.6.2, il deep learning costituisce lo strumento di analisi dati più potente di cui si disponga al giorno d'oggi e la sua applicazione in radiomica potrebbe acconsentire a notevoli sviluppi positivi. Di seguito si riporta uno studio del Dipartimento di Radiation Oncology dell'Harvard Medical School risalente al 2018 in cui vengono valutate le prestazioni del deep learning nella previsione del tasso di sopravvivenza, in una finestra di 2 anni dalla diagnosi, di pazienti affetti da NSCLC a partire da immagini di CT [38].

Questo paper permette di apprezzare molteplici modi in cui il DL si inserisce proficuamente tra gli strumenti usati in radiomica, in particolare toccando i seguenti punti:

- La costruzione di un modello inter-istituto
- Le prestazioni predittive del DL
- La comparazione con modelli basati su altre tecniche di ML
- L'analisi di una mappa di attivazione per la comprensione del processo di apprendimento
- La correlazione con aspetti biologici del cancro.

Lo studio fa uso di 7 dataset indipendenti provenienti da 5 istituzioni diverse che conta un totale di 1,194 pazienti per costruire un modello predittivo basato su una CNN. In particolare, i pazienti sono distinti in due gruppi sulla base del trattamento ricevuto, radioterapia e chirurgia, in modo da valutare separatamente le capacità predittive di due varianti del modello: per la radioterapia viene, infatti, usata la CNN vera e propria, mentre per la chirurgia si applica il *transfer learning* del primo ad un secondo modello.

Il vasto dataset iniziale è stato suddiviso in tre sottoinsiemi per le diverse fasi della costruzione della rete neurale e, parallelamente, il modello a trasferimento: il primo per il training, il secondo per il *fine tuning* ed il terzo per il testing. Infine, un ultimo data set è stato adoperato per lo *stability assesment*.

Lo stesso studio si pone due ulteriori obiettivi. Il primo di questi è il confronto della CNN con un modello di analisi radiomica basato su classificatore *radnom forest*, ovvero un approccio differente di ML dove l'estrazione delle features avviene separatamente (si ricorda che nel DL l'estrazione viene fatta dalla stessa NN che effettua poi l'analisi e la pattern recognition, in un ciclo chiuso di auto apprendimento).

Il secondo obiettivo è l'esplorazione delle basi biologiche dei descrittori fenotipici quantitativi definiti dalla CNN, con l'intento di verificare l'ipotesi radiomica trattata nel primo capitolo. Tale esplorazione viene fatta studiando la correlazione dei descrittori sopracitati con i risultati di un saggio GSEA (Gene Set Enrichment Analysis), ovvero un'analisi genomica volta a

riconoscere pattern di espressione genica, identificando i geni sovra-espressi, in un sottogruppo dei pazienti usati nello studio.

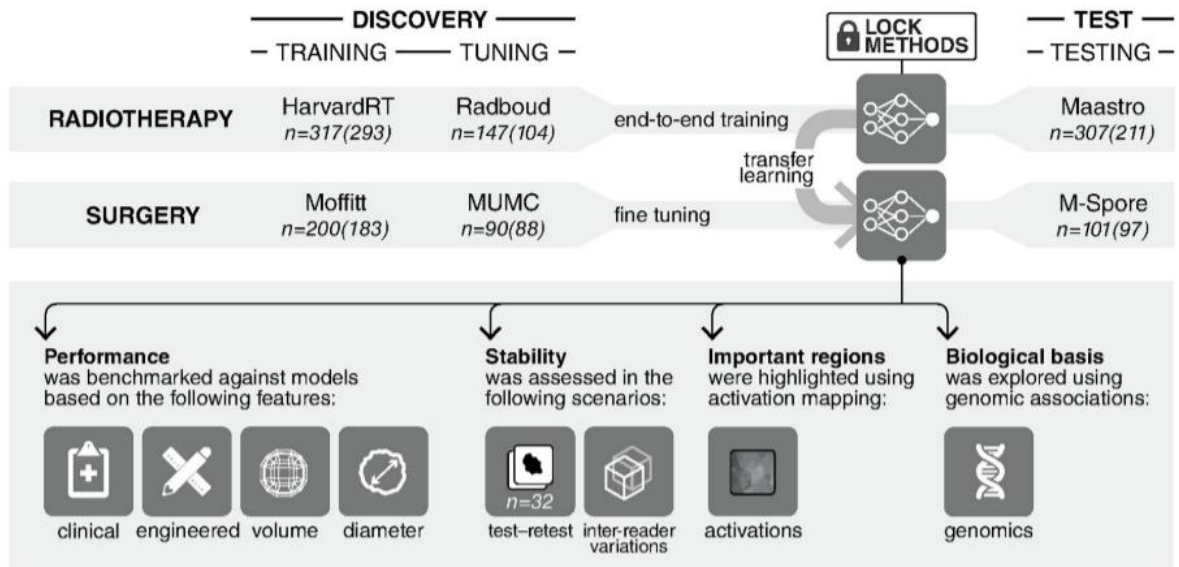


Figura 13: WorkFlow dello studio. Dallo schema si riconoscono le tre fasi di progettazione dei modelli di DL con riportati i nomi dei singoli dataset e la loro dimensione (n). Nella sezione inferiore sono riportate le diverse valutazioni che sono state fatte sui modelli.

L'architettura adottata è una CNN tridimensionale composta da 4 strati convoluzionali, l'ultimo di questi è usato per la generazione della mappa di attivazione con un metodo *gradient-weighted*. Le prestazioni predittive dell'overall survival dei modelli ottenuti sono state valute con il consueto calcolo della AUROC

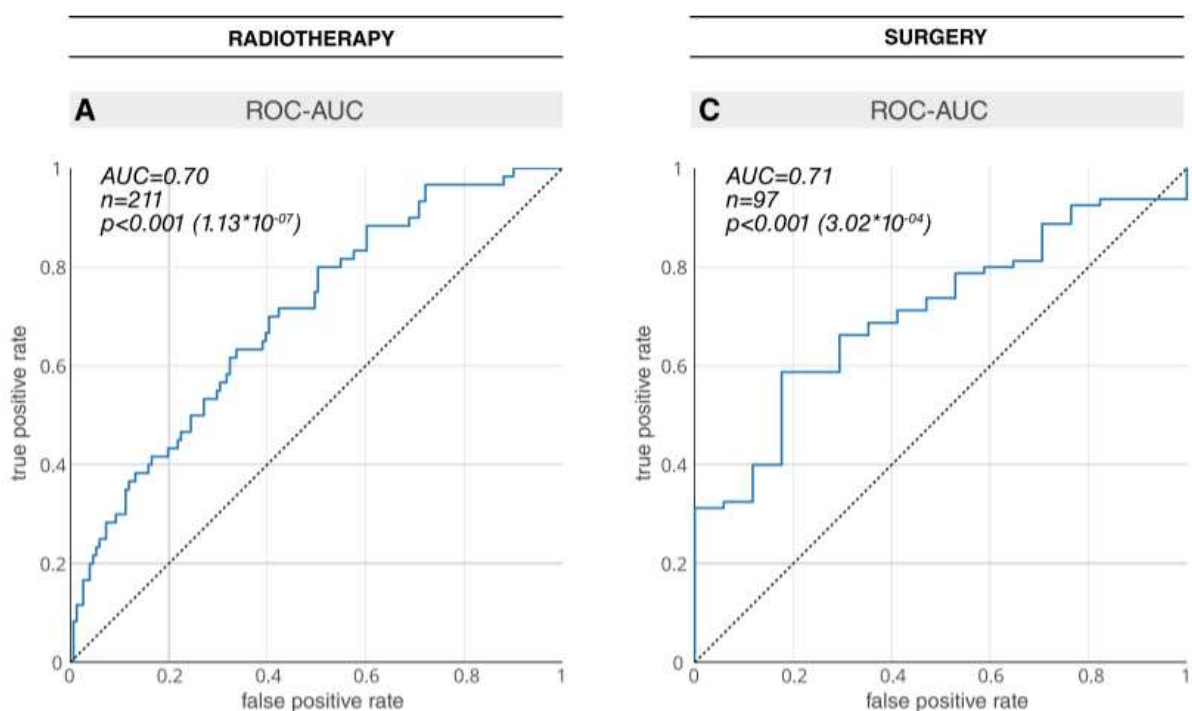


Figura 14: ROC del predittore CNN a sinistra e del modello transfer a destra.

e con una stima di Kaplan-Meier, di cui si riporta la curva di sopravvivenza.

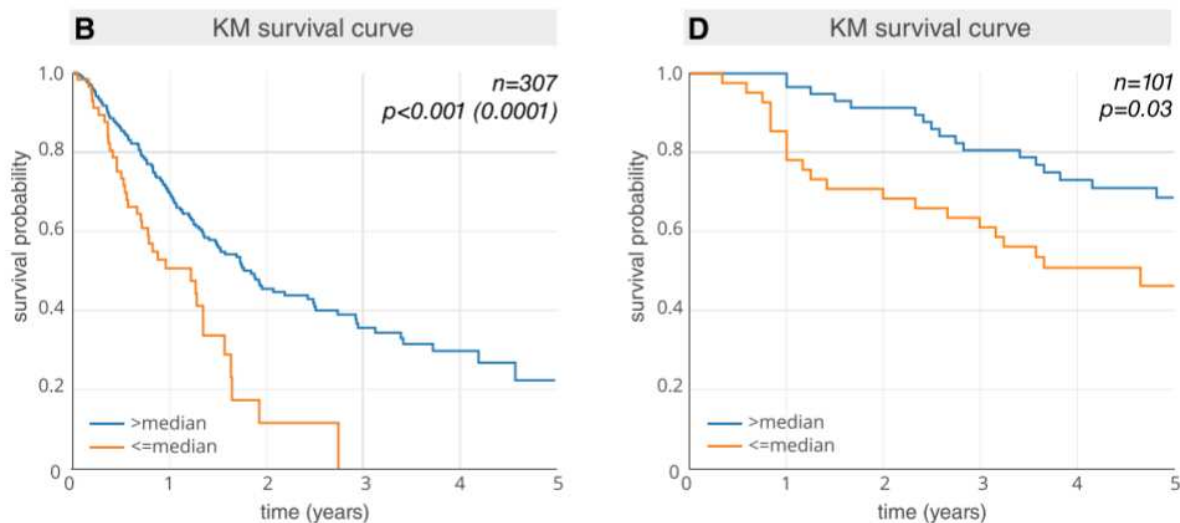


Figura 15: Il diagramma di Kaplan-Meier, dallo stimatore omonimo, riporta in ordinata la probabilità di sopravvivenza ed in ascissa l'asse temporale, approssimando con buona precisione la funzione di sopravvivenza effettiva del fenomeno studiato.

Dai grafici è possibile osservare l'abilità della NN nel predire l'OS e stratificare i pazienti nei gruppi di alto e basso rischio di mortalità.

L'analisi di attivazione gradient-weighted permette di evidenziare le regioni più rilevanti per la costruzione della rete e con il maggiore impatto sulle previsioni eseguite. I ricercatori hanno osservato che la rete tende a fissarsi sull'interfaccia tra tumore e stroma (parenchima o pleura). Riportano, poi, che la maggior parte dei contributi alle previsioni sono arrivati sotto forma di grandi aree ininterrotte di densità relativamente più elevata, esempi di questi includono aree dei lobi poco vascolarizzate o interfacce frastagliate tra aree con maggiore differenza di intensità. In aggiunta, è stato osservato che il tessuto normale, come il tessuto osseo ad alta densità, è sostanzialmente ignorato dalla NN, probabilmente perché ritenuto non informativo data la sua presenza nella maggioranza delle immagini.

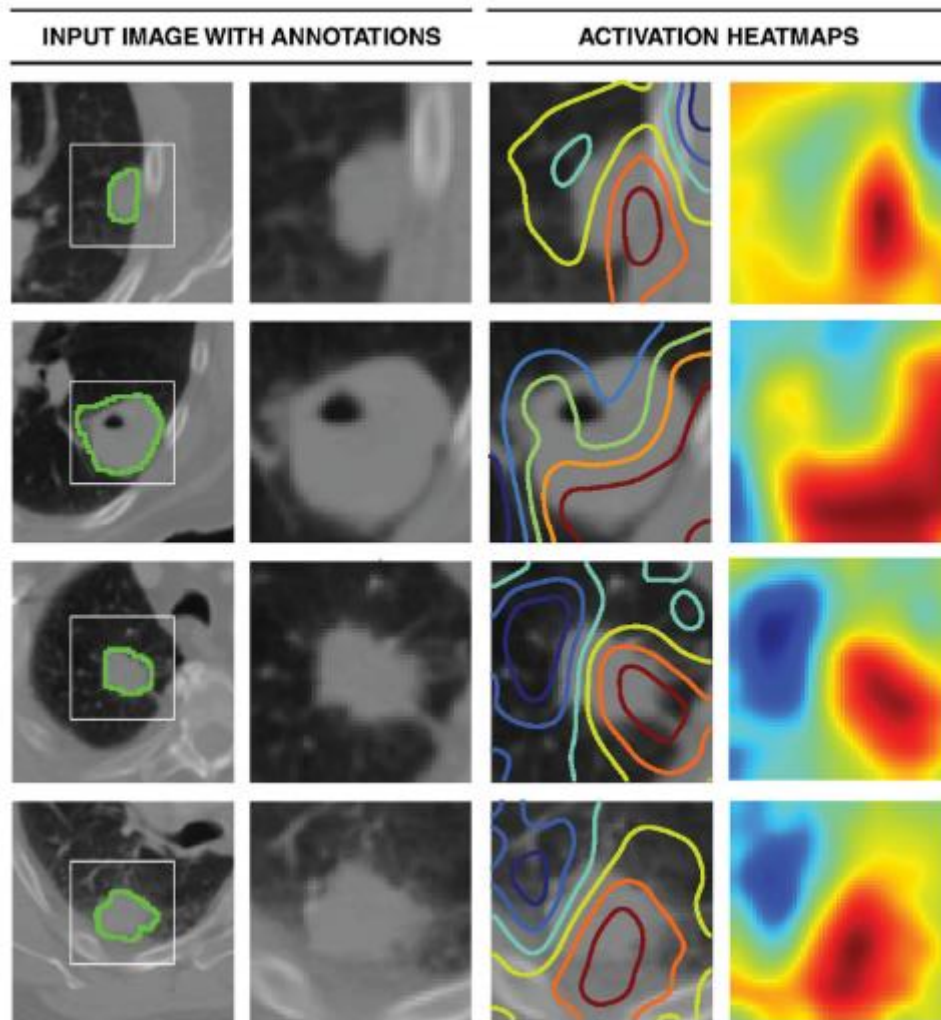


Figura 16: Sono riportati 4 esempi di attivazione con (da sinistra a destra): ROI, ritaglio della ROI, contorni di attivazione, heatmap di attivazione (crescente dal blu al rosso).

Per il confronto della CNN viene adottato un classificatore random forest (modello di ML non a rete neurale) applicato ad una raccolta di 40 features estratte dal set di immagini con l'uso di PyRadiomics e selezionate con il metodo mRMR (minimum redundancy maximum relevance). La selezione comprende descrittori di: forma del tumore, intensità dei voxel (istogramma) e textures del secondo ordine. Il modello ingegnerizzato (random forest) ha dimostrato una prestazione prognostica di $AUC = 0,66$ (IC al 95%) e $AUC = 0,58$ (IC al 95%) per, rispettivamente, i data-set di radioterapia e chirurgia. In generale, comparando questi numeri con i grafici in figura 14, si osserva che la rete di deep learning ha ottenuto performance migliori rispetto al modello RF per entrambi i sottogruppi di pazienti, ma questa differenza è stata nettamente più significativa per i pazienti sottoposti a intervento chirurgico. Nell'ultima analisi vengono studiate le correlazioni tra i descrittori quantitativi ottenuti con la CNN e i risultati di un'indagine genomica dei geni sovra-espresi nella stessa coorte. In

particolare, si riscontra che le features di DL più significative per la previsione dell'OS mostrano forte correlazione con i geni appartenenti ai *pathways* del ciclo cellulare e dei processi trascrizionali, suggerendo che la stessa estrazione delle features sia “guidata” da processi molecolari sottostanti legati alla proliferazione e progressione del tumore. Ad esempio, la sinapsi meiotica, l'impacchettamento dei telomeri e vari stadi del ciclo cellulare come G1 e S sono tra le principali associazioni ottenute, le quali sono di carattere negativo, denotando una correlazione inversa con la capacità proliferativa del fenotipo.

In conclusione, la CNN ha mostrato l'esistenza di importanti correlazioni tra le textures di un'immagine e le peculiarità genomiche del tumore, come espresso dall'ipotesi radiomica.

Studi come quello questo permettono quindi di apprezzare la potenza del deep learning nel contesto radiomico, identificando un'ovvia direzione in cui questa tecnologia si evolverà.

5. CONCLUSIONI

Nel primo capitolo è stato presentato un quadro generale dell'uso delle immagini in medicina, concentrandosi sulle immagini di radiologia, medicina nucleare, ecografia e risonanza magnetica, le quali sono ampiamente utilizzate e costituiscono tutt'oggi un elemento fondamentale per la diagnosi di diverse patologie. L'attenzione è stata poi posta sul campo oncologico, caratterizzato da aspetti distintivi riguardanti la formulazione di diagnosi, prognosi e trattamento. Successivamente, la breve illustrazione dei sistemi di assistenza alla diagnosi, i CAD, ha permesso di ricostruire il ruolo ricoperto dalle tecnologie informatiche nella pratica medica. Si è così potuta introdurre la radiomica all'interno del contesto clinico odierno, evidenziando la forte associazione che questa tecnica di analisi ha con aspetti caratteristici delle patologie tumorali e quindi con la costruzione di trattamenti personalizzati. A questo proposito alcuni degli studi riportati hanno fornito evidenze concrete a sostegno dell'ipotesi radiomica esposta nel primo capitolo sottolineando il ruolo che questa tecnica di analisi può ricoprire nella definizione dei biomarcatori, con i quali si rende possibile la terapia mirata individuale. Parallelamente, gli studi citati hanno dimostrato le ottime prestazioni dei modelli diagnostici e prognostici radiomici confermando le effettive potenzialità della radiomica come strumento di supporto alla decisione clinica. D'altro canto, nell'ultimo capitolo, sono stati messi in evidenza aspetti critici, emergenti in diverse fasi del flusso radiomico, che ne limitano fortemente la diffusione e l'espansione come tecnologia. Di questi, la standardizzazione è il problema principale, poiché nella medicina moderna, basata sull'evidenza scientifica, una determinata procedura, tecnica, farmaco o sostanza viene approvata solo se vi è sufficiente evidenza del suo funzionamento ed efficacia, anche in relazione ad eventuali effetti avversi. La mancanza di standard a più livelli (dalla configurazione dei sistemi di acquisizione alla costruzione dei modelli di IA) rende difficile effettuare studi inter-istituto e convalidare i risultati del singolo studio su scala più ampia. Allo stesso tempo vi è la concreta impossibilità di definire precisamente un metodo di lavoro condiviso adottabile dai singoli istituti.

Per questo motivo è necessario, da parte delle figure coinvolte, compiere uno sforzo collettivo di definizione di standard che permettano di consolidare le procedure. Un punto di partenza di tale sforzo è la costruzione di database condivisi su larga scala che superino i problemi legati alla privacy ed alla gestione dei dati e consentano di sfruttare a pieno le potenzialità del data mining, seguito dall'identificazione delle features di maggiore rilevanza clinica per la diagnosi oncologica. A questo scopo, come già mostrato nei paragrafi precedenti, sono state create diverse iniziative grazie a cui si stanno osservando i primi risultati delle collaborazioni nate e se ne stanno apprezzando i benefici.

Contemporaneamente la radiomica deve affrontare un miglioramento delle tecniche analitiche attraverso la costruzione e validazione di modelli basati sull'apprendimento automatico sempre più performanti. Nello specifico, in funzione del tipo di machine learning adottato, si distinguono le seguenti criticità: nei modelli supervisionati è richiesta la consolidazione di particolari tecniche ed algoritmi statistici per uniformare la pre-selezione dei dati da utilizzare nel training e garantire maggiore ripetibilità dei test; nel caso dei modelli non supervisionati invece, non essendo necessaria un'azione preventiva dell'operatore, vi è maggiore libertà nello sviluppo del modello, ma insorgono problematiche relative all'interpretazione dei risultati. Il ML non supervisionato ha infatti la capacità di riconoscere associazioni, correlazioni e legami di dipendenza tra i dati in modo autonomo ed originale; questo consente anche di ottenere osservazioni ed insights prima sconosciuti utili nella comprensione della patologia studiata e, talvolta, di produrre nuova conoscenza, ma porta con sé numerose difficoltà nell'interpretazione dei risultati che, come detto, sono propri della natura e peculiarità del modello usato.

Un'ultima utile osservazione possibile nell'uso dell'IA in radiomica riguarda la differente funzione che il deep learning ricopre nella fase finale di mining rispetto ai compiti di computer vision svolti, nella diagnostica per immagini, con la segmentazione. Le CNN sono ormai consolidate come modello altamente prestazionale nella visione artificiale, risultando adatte alla segmentazione dei tessuti. Nell'ultimo studio riportato una CNN è stata invece usata per l'analisi delle features a fini prognostici, assolvendo ad un compito fondamentalmente diverso: la ricerca di informazioni contenute nell'immagine che vanno oltre l'elaborazione visiva senza richiedere, quindi, di simulare la visione umana.

In conclusione, la radiomica può essere definita una tecnologia promettente, che nel vicino futuro potrà essere integrata nei sistemi di supporto alla diagnosi dei tumori al fianco delle altre tecniche di analisi omica, consentendo la formulazione di prognosi precise e trattamenti efficaci adattabili al singolo paziente ed alle mutazioni in cui la patologia incorre inevitabilmente. Inoltre, le sue potenzialità potranno essere usate anche in altri campi della medicina dove si fa uso delle bioimmagini, per la cura e diagnosi di patologie di altra natura difficili da studiare con gli usuali metodi di indagine clinica, come ad esempio quelle neurologiche.

BIBLIOGRAFIA

- [1] J. H. Scatliff e P. J. Morris, «From Röntgen to Magnetic Resonance Imaging: The History of Medical Imaging», 2014.
- [2] W. G. Bradley, «History of Medical Imaging», 2008.
- [3] G. Coppini, S. Diciotti, e G. Valli, *Bioimmagini, Collana di Ingegneria Biomedica*, III Edizione., vol. n° 8. Bologna: Pàtron Editore, 2012.
- [4] K. Doi, «Computer-aided diagnosis in medical imaging: Historical review, current status and future potential», *Computerized Medical Imaging and Graphics*, vol. 31, n. 4–5, pagg. 198–211, 2007, doi: 10.1016/j.compmedimag.2007.02.002.
- [5] R. L. Engle, «Attempts to Use Computers as Diagnostic Aids in Medical Decision Making: A Thirty-Year Experience», *Perspect Biol Med*, vol. 35, n. 2, pagg. 207–219, 1992, doi: 10.1353/pbm.1992.0011.
- [6] J. Yanase e E. Triantaphyllou, «A systematic survey of computer-aided diagnosis in medicine: Past and present developments», *Expert Systems with Applications*, vol. 138. Elsevier Ltd, 30 dicembre 2019. doi: 10.1016/j.eswa.2019.112821.
- [7] L. Fass, «Imaging and cancer: A review», *Molecular Oncology*, vol. 2, n. 2. Elsevier, pagg. 115–152, 2008. doi: 10.1016/j.molonc.2008.04.001.
- [8] I. Dagogo-Jack e A. T. Shaw, «Tumour heterogeneity and resistance to cancer therapies», *Nature Reviews Clinical Oncology*, vol. 15, n. 2. Nature Publishing Group, pagg. 81–94, 1 febbraio 2018. doi: 10.1038/nrclinonc.2017.166.
- [9] R. Fisher, L. Pusztai, e C. Swanton, «Cancer heterogeneity: Implications for targeted therapeutics», *British Journal of Cancer*, vol. 108, n. 3. pagg. 479–485, 19 febbraio 2013. doi: 10.1038/bjc.2012.581.
- [10] P. Lambin *et al.*, «Radiomics: Extracting more information from medical images using advanced feature analysis», *Eur J Cancer*, vol. 48, n. 4, pagg. 441–446, mar. 2012, doi: 10.1016/j.ejca.2011.11.036.
- [11] F. Shaikh, D. Dupont-Roettger, J. Dehmeshki, O. Awan, O. Kubassova, e S. Bisdas, «The Role of Imaging Biomarkers Derived From Advanced Imaging and Radiomics in the Management of Brain Tumors», *Front Oncol*, vol. 10, set. 2020, doi: 10.3389/fonc.2020.559946.
- [12] S. Volpe *et al.*, «Blood- and Imaging-Derived Biomarkers for Oncological Outcome Modelling in Oropharyngeal Cancer: Exploring the Low-Hanging Fruit», *Cancers (Basel)*, vol. 15, n. 7, apr. 2023, doi: 10.3390/cancers15072022.

- [13] R. J. Gillies, P. E. Kinahan, e H. Hricak, «Radiomics: Images are more than pictures, they are data», *Radiology*, vol. 278, n. 2, pagg. 563–577, feb. 2016, doi: 10.1148/radiol.2015151169.
- [14] R. T. H. M. Larue, G. Defraene, D. De Ruyscher, P. Lambin, e W. Van Elmpt, «Quantitative radiomics studies for tissue characterization: A review of technology and methodological procedures», *British Journal of Radiology*, vol. 90, n. 1070. British Institute of Radiology, 2017. doi: 10.1259/bjr.20160665.
- [15] F. Kiani, «Texture features in medical image analysis: a survey».
- [16] M. K. Ghalati, A. Nunes, H. Ferreira, P. Serranho, e R. Bernardes, «Texture Analysis and Its Applications in Biomedical Imaging: A Survey», *IEEE Reviews in Biomedical Engineering*, vol. 15. Institute of Electrical and Electronics Engineers Inc., pagg. 222–246, 2022. doi: 10.1109/RBME.2021.3115703.
- [17] M. A. Member Ieee e R. King, «Texural Features Corresponding to Texural Properties».
- [18] W. Ma *et al.*, «Breast Cancer Molecular Subtype Prediction by Mammographic Radiomic Features», *Acad Radiol*, vol. 26, n. 2, pagg. 196–201, feb. 2019, doi: 10.1016/j.acra.2018.01.023.
- [19] B. Koçak, E. Ş. Durmaz, E. Ateş, e Ö. Kılıçkesmez, «Radiomics with artificial intelligence: A practical guide for beginners», *Diagnostic and Interventional Radiology*, vol. 25, n. 6. AVES, pagg. 485–495, 1 novembre 2019. doi: 10.5152/dir.2019.19321.
- [20] K. Yasaka e O. Abe, «Deep learning and artificial intelligence in radiology: Current applications and future directions», *PLoS Medicine*, vol. 15, n. 11. Public Library of Science, 1 novembre 2018. doi: 10.1371/journal.pmed.1002707.
- [21] B. Kayalibay, G. Jensen, e P. van der Smagt, «CNN-based Segmentation of Medical Imaging Data», gen. 2017, [Online]. Disponibile su: <http://arxiv.org/abs/1701.03056>
- [22] M. W. Wagner, K. Namdar, A. Biswas, S. Monah, F. Khalvati, e B. B. Ertl-Wagner, «Radiomics, machine learning, and artificial intelligence—what the neuroradiologist needs to know», *Neuroradiology*, vol. 63, n. 12. Springer Science and Business Media Deutschland GmbH, pagg. 1957–1967, 1 dicembre 2021. doi: 10.1007/s00234-021-02813-9.
- [23] L. Nicosia *et al.*, «Prediction of the Malignancy of a Breast Lesion Detected on Breast Ultrasound: Radiomics Applied to Clinical Practice», *Cancers (Basel)*, vol. 15, n. 3, feb. 2023, doi: 10.3390/cancers15030964.
- [24] Z. Li *et al.*, «Diagnostic Performance of Mammographic Texture Analysis in the Differential Diagnosis of Benign and Malignant Breast Tumors», *Clin Breast Cancer*, vol. 18, n. 4, pagg. e621–e627, ago. 2018, doi: 10.1016/j.clbc.2017.11.004.

- [25] F. Botta *et al.*, «Association of a CT-based clinical and radiomics score of non-small cell lung cancer (NSCLC) with lymph node status and overall survival», *Cancers (Basel)*, vol. 12, n. 6, giu. 2020, doi: 10.3390/cancers12061432.
- [26] S. Rizzo, F. Savoldi, D. Rossi, e M. Bellomi, «Radiogenomics as association between non-invasive imaging features and molecular genomics of lung cancer», *Ann Transl Med*, vol. 6, n. 23, pagg. 447–447, dic. 2018, doi: 10.21037/atm.2018.11.17.
- [27] S. Volpe, F. Mastroleo, M. Krengli, e B. A. Jereczek-Fossa, «Quo vadis Radiomics? Bibliometric analysis of 10-year Radiomics journey», *Eur Radiol*, 2023, doi: 10.1007/s00330-023-09645-6.
- [28] A. Savini *et al.*, «The role of acquisition angle in digital breast tomosynthesis: A texture analysis study», *Applied Sciences (Switzerland)*, vol. 10, n. 17, set. 2020, doi: 10.3390/app10176047.
- [29] G. Carloni *et al.*, «Brain metastases from NSCLC treated with stereotactic radiotherapy: prediction mismatch between two different radiomic platforms», *Radiotherapy and Oncology*, vol. 178, gen. 2023, doi: 10.1016/j.radonc.2022.11.013.
- [30] R. Ortiz-Ramon, A. Larroza, E. Arana, e D. Moratal, «A radiomics evaluation of 2D and 3D MRI texture features to classify brain metastases from lung cancer and melanoma», in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, lug. 2017, pagg. 493–496. doi: 10.1109/EMBC.2017.8036869.
- [31] Z. Qian *et al.*, «Differentiation of glioblastoma from solitary brain metastases using radiomic machine-learning classifiers», *Cancer Lett*, vol. 451, pagg. 128–135, giu. 2019, doi: 10.1016/j.canlet.2019.02.054.
- [32] Y. J. CHA *et al.*, «Prediction of Response to Stereotactic Radiosurgery for Brain Metastases Using Convolutional Neural Networks», *Anticancer Res*, vol. 38, n. 9, pagg. 5437–5445, set. 2018, doi: 10.21873/anticancer.12875.
- [33] N. Eklund *et al.*, «Extending the Minimum Information About Biobank Data Sharing Terminology to Describe Samples, Sample Donors, and Events», *Biopreserv Biobank*, vol. 18, n. 3, pagg. 155–164, giu. 2020, doi: 10.1089/bio.2019.0129.
- [34] A. Zwanenburg *et al.*, «The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping», *Radiology*, vol. 295, n. 2, pagg. 328–338, mag. 2020, doi: 10.1148/radiol.2020191145.
- [35] B. Kocak, B. Baessler, R. Cuocolo, N. Mercaldo, e D. Pinto dos Santos, «Trends and statistics of artificial intelligence and radiomics research in Radiology, Nuclear

- Medicine, and Medical Imaging: bibliometric analysis», *Eur Radiol*, giu. 2023, doi: 10.1007/s00330-023-09772-0.
- [36] N. Garau *et al.*, «Integrating Biological and Radiological Data in a Structured Repository: a Data Model Applied to the COSMOS Case Study», *J Digit Imaging*, 2022, doi: 10.1007/s10278-022-00615-w.
- [37] V. P. Balachandran, M. Gonen, J. J. Smith, e R. P. DeMatteo, «Nomograms in oncology: more than meets the eye», *Lancet Oncol*, vol. 16, n. 4, pagg. e173–e180, apr. 2015, doi: 10.1016/S1470-2045(14)71116-7.
- [38] A. Hosny *et al.*, «Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study», *PLoS Med*, vol. 15, n. 11, nov. 2018, doi: 10.1371/journal.pmed.1002711.