UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA BIOMEDICA

"DEEP LEARNING E IMMAGINI ISTOPATOLOGICHE: VERSO LA DIAGNOSI AUTOMATIZZATA"

Relatore: Prof. Augusto Ferrante

Laureando/a: Giulia Macaluso

ANNO ACCADEMICO 2024 - 2025

Data di laurea 26 settembre 2025

Indice

Abstract 2

Introduzione 3

1 Istopatologia digitale: immagini e diagnosi clinica 5

1.1 Cos'è l'istopatologia 5

1.2 Dal metodo tradizionale alla digitalizzazione 5

1.3 Immagini istopatologiche 6

1.3.1 Whole Slide Imaging (WSI) 7

1.3.2 Pre-elaborazione delle immagini 8

2 Fondamenti di intelligenza artificiale e reti neurali 10

2.1 Definizioni: AI, machine learning e deep learning. 10

2.2 Machine learning e caratteristiche handcrafted 11

2.3 Reti neurali 12

2.4 Reti neurali convoluzionali (CNN) 14

2.4.1 Feature extraction e classificazione di immagini 14

2.4.2 Architettura base e matematica 15

3 Classificazione automatica di immagini istopatologiche 19

3.1 Etichettatura e dataset 19

3.2 Training e modalità di apprendimento 20

3.2.1 Apprendimento supervisionato e non-supervisionato 20

3.2.2 Learning from scratch e transfer learning 21

3.3 Metodi di classificazione 22

3.4 Validazione e metriche di valutazione 22

3.5 Test 25

4 Vantaggi, limiti e futuro dei sistemi di deep learning nella diagnostica 27

4.1 Vantaggi 27

4.2 Limiti e sfide attuali 28

4.3 Interpretazione dei risultati 30

4.4 Validazione clinica 32

4.5 Questioni etiche e regolatorie 33

4.6 Prospettive future 34

Conclusione 36

Sitografia 38

Elenco delle figure 40

Abstract

L'istopatologia, ovvero l'analisi microscopica di tessuti biologici, è uno degli strumenti principali per la diagnosi di patologie. Questa analisi si è basata in passato sulla sola valutazione manuale delle immagini, processo lungo e soggettivo, influenzato da variabilità dell'osservatore e soggetto a possibili errori umani. Negli ultimi anni, l'intelligenza artificiale ha portato a diverse innovazioni nell'ambito della patologia digitale. Sono stati sviluppati algoritmi di deep learning per l'elaborazione e classificazione automatica di immagini istologiche digitalizzate tramite sistemi di *Whole Slide Imaging* (WSI). Tali sistemi decisionali automatici possono affiancare il medico, per una diagnosi più rapida, accurata e imparziale.

Questa tesi introduce alcuni concetti fondamentali quali l'istopatologia e la sua digitalizzazione, machine learning, deep learning e reti neurali, oltre al ruolo centrale delle immagini WSI nei modelli predittivi. Vengono poi approfondite le reti neurali convoluzionali (CNN) nell'ambito della diagnostica istopatologica automatizzata, dalla fase di training, all'elaborazione e classificazione delle immagini. Sono inoltre valutati i principali dataset disponibili e l'interpretabilità degli algoritmi. La tesi si conclude esponendo vantaggi, limiti e prospettive future dell'intelligenza artificiale nell'ambito della diagnostica istopatologica.

Introduzione

La patologia digitale si sta rivelando un'innovazione in continuo cambiamento e sempre più rilevante in ambito della medicina moderna. Il progresso in questo campo è stato favorito dalla digitalizzazione dei vetrini istologici, in particolare dall'utilizzo di tecnologie di *Whole Slide Imaging*, per la creazione di ampi archivi di immagini digitali di alta qualità. Una svolta fondamentale nella patologia digitale ricade sull'integrazione di sistemi basati sull'intelligenza artificiale. I nuovi algoritmi di deep learning hanno un grande potenziale per l'analisi di immagini istopatologiche, in quanto sono in grado di classificare i tessuti e rilevare metastasi, dimostrandosi utili strumenti di supporto nella diagnosi e nella prognosi. Tuttavia, ci sono ancora ostacoli tecnici, etici e regolatori, oltre che un certo scetticismo, per l'implementazione di questi sistemi nella reale pratica clinica.

L'obiettivo di questa tesi è quello di esplorare le basi e i processi fondamentali del deep learning in istopatologia digitale, evidenziando vantaggi e limiti dei sistemi di diagnosi automatica.

Per iniziare, nel primo capitolo viene introdotto il contesto generale della patologia digitale, partendo dalla pratica della diagnosi tradizionale e la sua evoluzione verso la digitalizzazione; viene posta una certa attenzione sulle nuove tecnologie per l'acquisizione di immagini istopatologiche, in particolare i processi di *Whole Slide Imaging* e quelli di pre-elaborazione delle immagini.

Il capitolo 2, fornisce dei concetti fondamentali per comprendere meglio i sistemi basati sull'intelligenza artificiale. Vengono, innanzitutto, chiarite le definizioni di AI, machine learning e deep learning; dopodiché sono presentati il metodo di apprendimento basato su caratteristiche *handcrafted*, il concetto di reti neurali e il funzionamento di una rete neurale convoluzionale profonda, tra i modelli più utilizzati ed efficaci per l'elaborazione delle immagini.

Nel capitolo 3, vengono approfonditi i processi di classificazione automatica delle immagini istopatologiche: partendo dalla costruzione e l'etichettatura di dataset, passando poi alle diverse strategie di addestramento dei modelli e di classificazione, per poi concludere con le metriche utilizzate nelle fasi di validazione e test dei sistemi.

Nel capitolo 4, si analizzano in maniera critica i vantaggi, i limiti e le prospettive future dei sistemi di deep learning per la diagnosi. Particolare attenzione è riposta sul tema della mancanza di interpretabilità dei risultati, presentando alcuni metodi sviluppati per risolvere il problema dei cosiddetti modelli di deep learning *black-box*. Si discutono poi le sfide nella validazione

clinica, necessaria per l'adozione dei sistemi AI nella pratica medica, oltre alle questioni etiche e regolatorie.

La tesi si conclude con una breve riflessione sulle potenzialità del deep learning in medicina, come alleato strategico del patologo, capace di ottimizzare risorse e migliorare il processo di diagnosi, contribuendo al progresso della medicina personalizzata e di precisione.

Capitolo 1
Istopatologia digitale: immagini e diagnosi clinica

In questo capitolo vengono trattati alcuni concetti fondamentali legati all'anatomia patologica e alla diagnosi. In primo luogo, si definisce cos'è l'istopatologia. Si illustra come i metodi tradizionali dell'istologia sono evoluti in quella che è la patologia digitale, evidenziando le principali differenze. Infine si esplora il ruolo essenziale delle immagini istopatologiche digitali, utilizzabili nei sistemi di intelligenza artificiale per la diagnosi.

1.1 Cos'è l'istopatologia

L'anatomia patologica studia le alterazioni macroscopiche e microscopiche dovute a stati patologici in organi e tessuti umani, che mira a fornire una diagnosi dall'interpretazione delle alterazioni [1].

L'istopatologia è una *"parte dell'anatomia patologica che studia le alterazioni strutturali dei tessuti con i metodi dell'istologia"* [2]. Tramite biopsia è possibile distinguere il cancro dalle cellule sane, oltre che i tumori maligni da quelli benigni, in modo da poter scegliere il giusto approccio terapeutico e clinico.

Prendendo come esempio il cancro, una tra le principali cause di morte, è di fondamentale importanza combatterlo partendo proprio da una diagnosi tempestiva tramite metodi di rilevazione efficienti.

1.2 Dal metodo tradizionale alla digitalizzazione

Tradizionalmente, la diagnosi istologica da parte degli esperti patologi si basa su diverse componenti: individuazione di cellule alterate e caratteristiche morfologiche tramite riconoscimento visivo, misurazione di caratteristiche microscopiche rilevanti e classificazione dei tessuti osservati, confrontandoli con pattern patologici e dati clinici già documentati [3].

Le immagini dei tessuti vengono raccolte ed esaminate al microscopio, per individuare cellule cancerogene o elementi anomali, dai patologi. Questo processo manuale risulta essere soggettivo, qualitativo e condizionato da errori umani. Inoltre, essendo oggigiorno passati a procedure non invasive per l'acquisizione dei campioni tissutali, l'analisi ed identificazione risulta essere più difficile per il patologo, a causa delle ridotte dimensioni e qualità dei campioni.

Negli ultimi anni, il campo della patologia ha fatto un passo avanti verso l'automatizzazione, attraverso l'integrazione di tecnologie digitali nei processi diagnostici.
La patologia digitale (DP) è definita come la pratica della patologia che utilizza immagini digitali di sezioni microscopiche totali, tecniche avanzate di *slide scanning* e algoritmi AI per rilevamento, segmentazione, diagnosi e analisi di immagini [4, 5].
Lo sviluppo di sistemi di *Whole Slide Imaging* (WSI) è stato il principale promotore della patologia digitale negli ultimi anni, che ha permesso l'acquisizione, la valutazione e l'interpretazione di immagini ad alta risoluzione di intere sezioni istologiche tramite scanner. Dando WSI come input, gli algoritmi sono in grado di trovare connessioni tra caratteristiche rilevate e campioni precedentemente etichettati, portando alla diagnosi e alla prognosi.
Questo processo di diagnosi digitalizzato presenta evidenti miglioramenti rispetto il metodo patologico tradizionale, in quanto è più oggettivo, quantitativo, accurato e costante. È basato su dati e permette la standardizzazione di criteri di valutazione dei tumori (ad esempio il punteggio di Gleason per la classificazione del carcinoma prostatico). Inoltre le immagini digitali sono più semplici da manovrare e condividere dei vetrini istologici.
Indubbiamente, la patologia digitale si espanderà in futuro, nonostante presenti ancora alcune limitazioni legate a problemi di campionamento (qualità e velocità), memoria di archiviazione delle immagini, standardizzazione, riluttanza dei patologi e necessità di molti dati annotati per database di immagini.
Ulteriore componente della patologia digitale sono i sistemi di intelligenza artificiale per categorizzare i tessuti, oltre a predire una prognosi e una risposta terapeutica, basandosi su caratteristiche istologiche [6].
In particolare, l'intelligenza artificiale usata in DP procura informazioni difficili da identificare per un patologo, quali l'analisi oggettiva di biomarcatori, la quantificazione di cellule, espressione, densità e schemi di distribuzione.
Secondo vari studi, l'accuratezza della patologia digitale con sola AI è comparabile (se non superiore) a quella di esperti patologi, essa migliora ulteriormente quando i sistemi AI vengono integrati nei normali protocolli eseguiti dall'uomo [5]. L'obiettivo principale dell'introduzione dell'intelligenza artificiale in patologia è infatti quello di supportare il patologo nella diagnosi, in modo da avere risultati migliori, più efficienti e consistenti.

1.3 Immagini istopatologiche

Alla base dell'istopatologia si trova l'analisi delle immagini dei tessuti. Tradizionalmente l'analisi patologica veniva eseguita su vetrini istologici tramite microscopio, mentre negli

ultimi anni si è passati a una digitalizzazione di questi vetrini grazie alla diffusione di scanner ad alta qualità che permettono un'analisi dettagliata al computer.

1.3.1 Whole Slide Imaging (WSI)

La tecnologia di *Whole Slide Imaging* consente la digitalizzazione completa dei vetrini istologici, generando immagini ad alta risoluzione che catturano anche i minimi dettagli.
Questi sistemi WSI includono due fasi: acquisizione delle immagini tramite scanner e analisi al computer.
La prima fase consiste nella scansione completa ad alta velocità e alta risoluzione del campione istologico e creazione del cosiddetto vetrino digitale. Questo file unico digitale viene generato da uno scanner WSI, che consiste in un microscopio unito a sistemi robotici e computerizzati.
Esistono diversi modelli di scanner e software in commercio, alcuni nomi includono: *Panoramic Scan II*, *Lamina*, *Pathscope* e *ScanScope*. Questi modelli hanno caratteristiche differenti, tra cui: la capacità di caricamento (pochi o molti vetrini contemporaneamente), la velocità di scansione e la modalità di acquisizione. I vetrini WSI sono preparati e colorati alla stessa maniera del metodo tradizionale al microscopio, ma vengono poi analizzati su monitor. I processi di preparazione delle immagini vengono meglio esplorati nel prossimo paragrafo.
Nella seconda fase, i vetrini digitali vengono esaminati al computer, tramite diversi software e algoritmi. Questa efficiente modalità permette di archiviare, analizzare, consultare e condividere rapidamente con altri esperti a distanza i file istologici [4].
L'analisi al computer di vetrini digitali WSI presenta parecchi vantaggi che includono:

- archiviazione efficiente e facilità di gestione dei file;
- accessibilità e velocità di condivisione;
- qualità dell'immagine e possibilità di esaminare la sezione di tessuto nel dettaglio, avendo la possibilità di ingrandire e ridurre l'immagine, annotare aree di interesse e misurare elementi rilevanti [7].

D'altra parte l'acquisizione dei minimi dettagli del tessuto con WSI comporta un problema di spazio di memoria, legato alla risoluzione elevata delle immagini.
Le *Whole Slide Images* trovano applicazione nella trasmissione di dati patologici per collaborazioni, standardizzazione e distribuzione di materiale per l'educazione e formazione di nuovi anatomopatologi e nell'analisi di immagini [4].
I vetrini istologici WSI possono inoltre essere forniti a sistemi di intelligenza artificiale per supportare il medico nella diagnosi del paziente. Questo può comportare un certo lavoro preliminare di etichettatura da parte del patologo, che deve annotare i file, associando ad ogni

immagine la classe di appartenenza (ad esempio tessuto sano o carcinogeno). Questo processo rientra nella fase di addestramento della rete, che verrà esposto in seguito.

1.3.2 Pre-elaborazione delle immagini.

La pre-elaborazione delle immagini è un processo necessario per convertire i dati in un formato che possa essere utilizzato come input per il software di elaborazione e per migliorarne certe caratteristiche. Utilizzare immagini istologiche che siano chiare e di alta qualità, è fondamentale per gli studi medici e la diagnosi, sia da parte di un patologo che di un algoritmo.

Questa fase preliminare include diverse tecniche che vanno a modificare dimensioni, orientamento, colore ed esposizione delle immagini.

- Ridimensionamento: consiste nella deformazione dell'immagine o nell'aggiunta di pixel per modificare le dimensioni del file, mantenendone le proporzioni. Questo metodo è utile per uniformare la dimensione dei campioni (caratteristica richiesta per alcuni strati delle CNN) e per concentrare l'analisi sulle zone di interesse. Ridurre le dimensioni delle immagini contribuisce inoltre a velocizzare il processo di addestramento dei modelli AI di diagnosi.
- Segmentazione: per separare il tessuto di interesse dalle zone non rilevanti, come lo sfondo.
- Normalizzazione del colore e dell'esposizione: il valore di intensità dei pixel viene normalizzato per assicurare consistenza e compensare differenze dei dati.
- Riduzione del rumore: vengono usati dei filtri per minimizzare distorsioni ed evidenziare le caratteristiche rilevanti per la diagnosi [8, 9].
- Compressione: per eliminare molti dei pixel ridondanti e superflui, poiché contribuiscono solo ad aumentare la complessità computazionale del modello, senza migliorare i risultati [10].

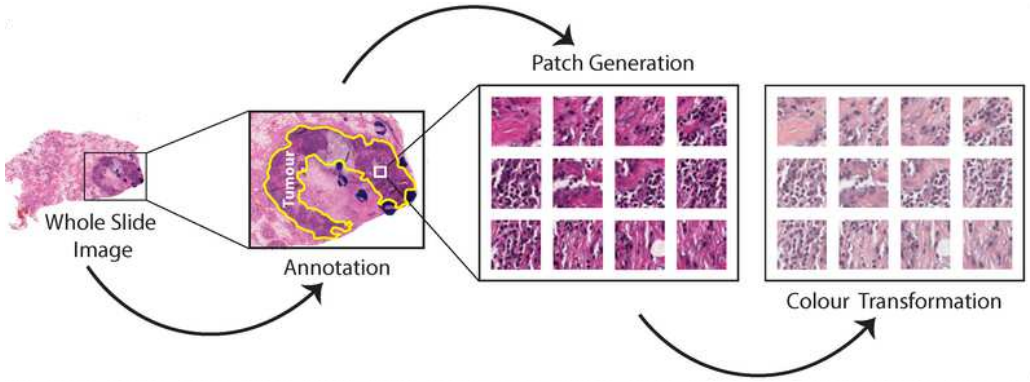

Figura 1: Esempio di modifica del colore nella fase di pre-elaborazione di sezioni di immagini istopatologiche digitali WSI.

Cambiamenti randomici, possono inoltre essere introdotti artificialmente alle immagini originali per il potenziamento dell'addestramento di software, in quanto aumentano la variabilità dei dati e migliorano la robustezza del modello. In questo modo, il sistema impara a riconoscere le variazioni da un più ampio range di input [11]. Questi cambiamenti includono: rotazioni, inversioni, modifiche della luminosità e dell'orientamento.

In conclusione la pre-elaborazione delle immagini istologiche è un processo essenziale per l'analisi diagnostica e l'addestramento di sistemi di intelligenza artificiale. L'obiettivo principale di questa prima fase è quello di ottimizzare la qualità e l'uniformità delle immagini, in modo da migliorare i modelli di analisi automatica, in termini di accuratezza e robustezza.

Capitolo 2
Fondamenti di intelligenza artificiale e reti neurali

L'intelligenza artificiale (AI) è già stata adotta negli ultimi anni in numerosi settori, compreso quello della patologia digitale ed è probabile che rivoluzionerà in un futuro prossimo i servizi sanitari e la diagnostica. Le tecniche di AI principalmente usate per l'elaborazione di immagini consistono nel machine learning (ML) e in particolare nel deep learning (DL). Questi sistemi automatici sono in grado di analizzare le *Whole Slide Images* per localizzare e classificare carcinomi e strutture anomale, oltre che esaminare la morfologia e l'organizzazione di elementi cellulari. Le loro prestazioni sono comparabili, se non superiori, a quelle di esperti patologi.
Per capire come l'AI venga utilizzato nella patologia digitale, è necessario definire alcuni concetti fondamentali. In questo capitolo, vengono innanzitutto esposte le differenze tra AI, machine learning e deep learning e le loro prime modalità di apprendimento. Dopodiché si analizzeranno le architetture più utilizzate, in particolare le reti neurali convoluzionali (CNN).

2.1 Definizioni: intelligenza artificiale, machine learning e deep learning

Il termine "intelligenza artificiale" è stato coniato per la prima volta dall'informatico statunitense McCarthy negli anni '50, riferendosi a tutte quelle macchine informatiche usate per provare a fare predizioni, ragionamenti e decisioni come l'uomo [5]. Un'altra possibile definizione più moderna è questa: *"L'intelligenza artificiale è una disciplina appartenente all'informatica che studia i fondamenti teorici, le metodologie e le tecniche che consentono la progettazione di sistemi hardware/software capaci di fornire prestazioni tipiche dell'intelligenza umana"* [7].
Gli approcci di machine learning, consistono negli algoritmi che permettono ad un sistema informatico di esaminare i dati in input e imparare automaticamente dall'esperienza passata per fare predizioni, adattando gli algoritmi a ogni nuova informazione. Possiamo altrimenti definire il ML come: *"capacità di una macchina di apprendere senza essere programmata esplicitamente"* [12].
Per finire, il deep learning, o apprendimento approfondito, è un particolare approccio di machine learning basato sulle reti neurali artificiali profonde, che simulano l'architettura delle reti neurali umane. Vengono definite profonde in quanto le reti neurali sono composte da molti strati, appunto modelli profondi con numerose fasi di elaborazione, le quali verranno trattate nei prossimi paragrafi. La particolarità degli algoritmi di deep learning, risiede nel fatto che

sono loro stessi a individuare le caratteristiche rilevanti dai numerosi dati in input per fornire una risposta, senza l'aiuto umano. Per questo motivo, l'apprendimento profondo è particolarmente adatto ai sistemi di visione artificiale (capacità di riconoscere elementi e pattern): in diagnostica è, infatti, la tecnica più usata per l'elaborazione e classificazione di immagini [7].

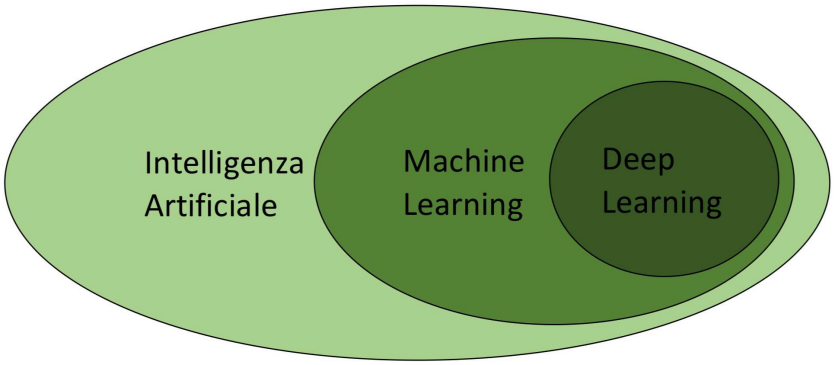

Figura 2: Diagramma rappresentativo della relazione tra intelligenza artificiale (AI), machine learning (ML) e deep learning (DL).

2.2 Machine learning e caratteristiche *handcrafted*

Nell'ambito dell'analisi delle immagini e dell'intelligenza artificiale, è importante distinguere due approcci. I primi sono quelli basati su caratteristiche progettate manualmente, i cosiddetti *handcrafted feautures approaches*, che utilizzano machine learning in una forma più semplice, e gli approcci più complessi di deep learning, basati su reti neurali profonde.

In questo paragrafo, viene spiegata velocemente la prima modalità di estrazione delle caratteristiche, mentre dal prossimo, l'attenzione sarà riposta sull'apprendimento profondo.

Gli algoritmi di ML si possono suddividere in due categorie, a seconda della conoscenza del dominio di applicazione e della specificità:

- *domain-agnostic handcrafted features approaches:* approcci basati su caratteristiche manuali indipendenti dal dominio. Utilizzano caratteristiche indipendenti dal contesto, da applicare a qualsiasi tipo di immagine (ad esempio colore, forme o texture generiche). Sono più generali e riutilizzabili in settori diversi senza grandi modifiche. Sono preferibili per esaminare immagini di malattie e tessuti di tipi diversi;
- *domain-specific handcrafted feautures approaches*: approcci basati su conoscenze specialistiche di un certo dominio, come l'istopatologia (ad esempio distribuzione delle cellule e rapporto nucleo/citoplasma). Sono metodi facilmente interpretabili dal patologo e molto specifici [5].

Questi metodi vengono usati per diagnosi, prognosi, classificazione di varie tipologie cancerogene, quando non è disponibile un gran numero di dati per l'addestramento; risultano essere più interpretabili e chiari per i patologi, rispetto ai metodi di DL [6].
La principale differenza tra questi approcci e il deep learning sta nel fatto che le caratteristiche rilevanti sono esplicitate dall'uomo sulla base della conoscenza del dominio, mentre gli algoritmi DL le apprendono automaticamente dal modello, senza assunzioni pre-esistenti.
I principali vantaggi dell'apprendimento profondo risiedono nel fatto che le reti neurali sono in grado di generare e studiare di conseguenza le caratteristiche più rilevanti per categorizzare le immagini. L'applicazione dei metodi DL per l'analisi di immagini patologiche, risulta più semplice e accurata; in generale migliore dell'approccio con ML [6].

2.3 Reti neurali

Una rete neurale è un tipo di algoritmo di machine learning, un modello computazionale, la cui architettura si ispira al funzionamento del cervello umano per prendere decisioni, *"utilizzando processi che imitano il modo in cui i neuroni biologici lavorano insieme per identificare fenomeni, valutare opzioni e trarre conclusioni*" [13].
Il modello di rete neurale artificiale più antico e semplice (un singolo neurone artificiale) è chiamato percettrone, ideato da Frank Rosenblatt nel 1958. Questo percettrone è in grado di risolvere problemi di classificazione lineare, ma è limitato, poiché non può apprendere funzioni non lineari complesse. Tali limiti sono stati superati con lo sviluppo dei percettroni multistrato (MLP), con più livelli di neuroni e funzioni di attivazione non lineari; questi sistemi stanno alla base delle moderne reti neurali profonde [13].
Una rete neurale è formata da neuroni artificiali, o nodi, organizzati in strati, o livelli (layers), connessi tra loro:

- livello di input: riceve i dati in entrata;
- livelli nascosti: possono essere uno o più strati, che modificano ed elaborano i dati ricevuti con pesi, bias e funzioni di attivazione. Più *hidden layers* ci sono, più è profonda e complessa la rete;
- livello di output: genera il risultato finale.

Ogni nodo può essere rappresentato da una relazione lineare tra input e output, con i seguenti elementi principali:

- pesi (w): determinano l'importanza relativa di ogni input (x);
- bias (b): valore costante, permette di spostare la funzione di attivazione;

- funzione di attivazione (f(x)): determina l'uscita (y), stabilisce se il neurone viene attivato. La formula generale è la seguente:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{1}$$

Il flusso di informazioni in una rete neurale funziona nel seguente modo: ogni input viene moltiplicato per il peso corrispondente, questi vengono poi sommati tra loro e viene aggiunto il valore di bias. Il risultato passa attraverso la funzione di attivazione, che lo trasforma in un output numerico. Questo valore, indipendentemente dalla sua intensità, viene trasmesso dal neurone artificiale al livello successivo, anche se in certi casi può risultare molto piccolo.

Le reti in cui il movimento di dati avviene in un'unica direzione, poiché l'uscita di un nodo attivo diventa l'ingresso del neurone successivo, sono dette di tipo *feedforward*.

Queste reti, comuni nei sistemi di deep learning, vengono addestrate grazie alla *backpropagation*, tecnica di retropropagazione, che opera in verso opposto (dall'output all'input), con lo scopo di ottimizzare i pesi e minimizzare progressivamente la funzione di errore del modello [13].

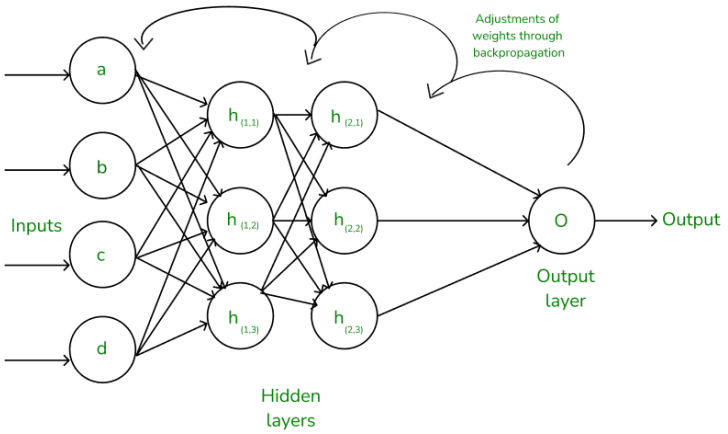

Figura 3: Rete neurale profonda, in cui sono evidenziati input, livelli nascosti, output e l'aggiornamento dei pesi attraverso retropropagazione dell'errore.

L'addestramento delle reti neurali richiede un gran numero di dati, in modo che il modello migliori progressivamente e automaticamente con l'esperienza. Dopodiché, i sistemi sono in grado di eseguire compiti complessi, come per esempio la classificazione di immagini, molto più velocemente dell'uomo. Le reti neurali sono alla base dei sistemi di deep learning usati per la diagnostica automatica e l'analisi istopatologica.

2.4 Reti neurali convoluzionali (CNN)

Una rete neurale convoluzionale (CNN) è un metodo di machine learning, comunemente applicato per il riconoscimento e la classificazione delle immagini e il più usato in patologia. È una classe di reti neurali artificiali *feedforward* profonde, che sfrutta i principi dell'algebra lineare (come moltiplicazione di matrice), per trovare modelli e strutture rilevanti in un'immagine. Queste reti sono formate da molti livelli sequenziali che possono calcolare un output dall'informazione in ingresso, smontando l'immagine in livelli gerarchici più piccoli e generando mappe delle caratteristiche (*feature map*) che descrivono le informazioni visive [6]. La particolarità delle CNN è la capacità di estrarre automaticamente le caratteristiche rilevanti, tramite operazioni di convoluzione, al contrario degli approcci con *handcrafted features* spiegati precedentemente.

Inoltre, una CNN richiede meno parametri, rispetto ad altri tipi di reti neurali, poiché la rete è invariante rispetto la traslazione dell'input, cioè non richiede che il modello impari diversi metodi di identificazione per uno stesso oggetto che si trova in punti diversi dell'immagine.

Alcuni esempi di architetture CNN più usate in patologia digitale sono le seguenti: *LeNet* di Yan LeCun (2 strati convoluzionali) e *AlexNet* di Krizhevsky (5 strati) sono reti meno profonde; *VGG19* (19 strati), *GoogLeNet* o *Inception* (22 strati) e *ResNet-50* sono reti molto più profonde e complesse, che riducono i parametri necessari e migliorano l'efficienza in fase di addestramento [14, 15].

2.4.1 Feature extraction e classificazione di immagini

Dopo lo step preliminare della pre-elaborazione delle immagini, le CNN svolgono due compiti principali: la *feature extraction*, ovvero l'estrazione delle caratteristiche rilevanti e la classificazione delle immagini.

- *Feature extraction*: questa prima fase consiste nell'acquisizione implicita delle caratteristiche da parte delle reti neurali convoluzionali. I primi strati superficiali apprendono caratteristiche semplici (come bordi e forme), mentre i livelli più profondi, quelle più complesse (come pattern e organizzazioni cellulari). Nonostante le somiglianze, l'apprendimento dei dettagli avviene diversamente per il patologo e per le macchine: mentre l'uomo si focalizza di più sulle forme, i computer si concentrano principalmente su patterns e strutture. In queste prime fasi, gli algoritmi analizzano automaticamente nuclei, citoplasma, strutture cellulari e aggiornano automaticamente i valori dei filtri durante la fase di addestramento, in modo da poter estrarre le caratteristiche computazionali rilevanti e predire finalmente un output.

- Classificazione: questo step utilizza una funzione di attivazione nello strato completamente connesso finale (*fully connected layer*) per inserire l'immagine nella corretta categoria ed etichettarla, basandosi sulle mappe delle caratteristiche estratte precedentemente [10].

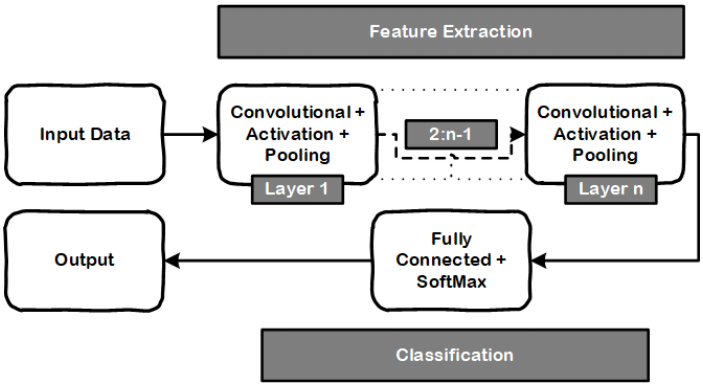

Figura 4: Schema a blocchi che evidenzia i due ruoli principali svolti nei livelli delle CNN: estrazione delle caratteristiche e classificazione.

2.4.2 Architettura base e matematica

La struttura delle CNN consiste in più livelli specializzati che permettono di estrarre caratteristiche, semplificarle e classificare l'immagine in input.

In questo paragrafo sono definiti i tre principali tipi di livelli: strati di convoluzione con funzione di attivazione, strati di *pooling* (questi due sono i cosiddetti livelli nascosti) e infine gli strati completamente connessi.

- Strati di convoluzione: usati per *feature extraction* automatica. L'operazione di convoluzione consiste nel far scorrere un filtro, detto *kernel* (una funzione lineare), sull'input da sinistra a destra. Di solito, un filtro è una matrice abbastanza piccola (ad esempio 3x3, 5x5) che esegue la convoluzione per individuare pattern locali. Matematicamente, avviene un prodotto scalare tra i valori dei pixel dell'immagine e i pesi del filtro in ogni posizione, è quindi una somma pesata tra il *kernel* e una porzione dell'immagine. Poiché il filtro scorre sull'immagine, questa operazione di convoluzione viene eseguita sull'intero input e genera come output una *feature map* (mappa delle caratteristiche o di attivazione), cioè una nuova matrice in cui ogni valore indica quanto una determinata caratteristica è presente in un punto preciso dell'immagine in entrata. Questo efficiente meccanismo di scorrimento permette quindi di usare un singolo filtro per elaborare una proprietà nell'intera immagine, indipendentemente dalla posizione, riducendone i parametri.

La formula che rappresenta l'operazione è la seguente:

$$F[i,j] = (I * K)_{[i,j]} = \sum_m \sum_n I_{[i+m,j+n]} \cdot K_{[m,n]} \tag{2}$$

In cui: F[i, j] è la *feature map* in uscita (in posizione i, j), I è l'immagine in ingresso di cui viene fatta la convoluzione ($*$) con il filtro K (con coordinate m, n).

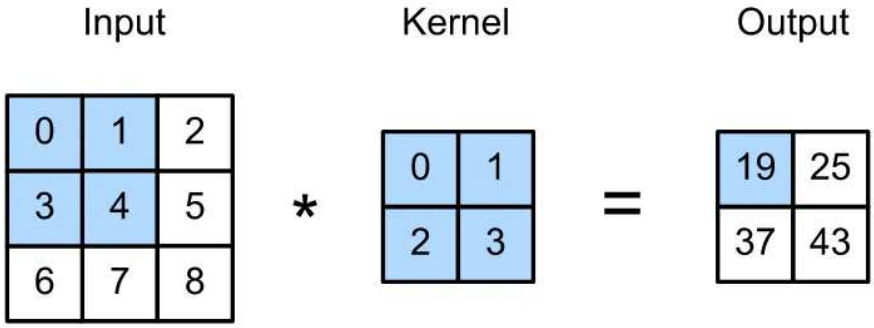

Figura 5: Operazione di convoluzione, tramite prodotto scalare tra valori dei pixel dell'immagine in input e filtro (*kernel*), che genera come output una *feature map*.

Le seguenti proprietà influiscono direttamente sulle dimensioni della *feature map*:

- *Stride*: è il passo con cui il filtro si muove sull'input, ovvero di quanti pixel il *kernel* si sposta in verticale e in orizzontale ad ogni step. Più grande è lo *stride*, minore è la risoluzione della mappa di attivazione, in quanto il filtro salta dei pixel, permettendo però di catturare le caratteristiche globali con meno parametri e meno memoria.
- *Padding*: è un meccanismo di compensazione, in cui viene aggiunta almeno una riga o colonna di pixel con valore zero a ogni lato dell'immagine in ingresso, prima della convoluzione. Gli obiettivi principali sono quelli di controllare le dimensioni della matrice in uscita e di elaborare meglio le caratteristiche ai bordi (che verrebbero meno esaminate rispetto a quelle centrali). Aggiungere troppi zeri, aumenta però la complessità e diminuisce l'efficienza.

Ogni dimensione dell'output, si può calcolare con questa equazione:

$$d_{out} = \frac{d_{in} + 2P - F}{S} + 1 \tag{3}$$

In cui: d è la dimensione presa in considerazione (larghezza, lunghezza) dell'output e dell'input, P è la quantità di zeri aggiunti a ogni lato, F la dimensione del filtro e S lo *stride*. Ad ogni livello convoluzionale possono essere applicati più filtri contemporaneamente. Ogni filtro cattura caratteristiche differenti e produce una *feature map* specifica; queste sono poi sovrapposte a formare un volume tridimensionale (in uscita dallo strato), la cui

profondità è data dal numero totale di *kernel*. È per questo che reti neurali più profonde sono in grado di elaborare immagini e caratteristiche più complesse [11, 16].

- Funzione di attivazione: l'output, proveniente dallo strato di convoluzione, viene passato attraverso una funzione di attivazione non lineare, in modo che il modello possa esprimere relazioni più complesse. La funzione più comune è *ReLU* (*Rectified Linear Unit*), che azzera i valori negativi e mantiene invariati quelli positivi.

$$R(x) = max(0, x) \tag{4}$$

- Strati di *pooling*: servono a ridurre la dimensione delle *feature map* attivate (quindi a diminuire il numero di parametri), mantenendo le informazioni più importanti, rendendo la rete più robusta a rumori e traslazioni e migliorando l'efficienza computazionale.

  Lo strato di *pooling* consiste in un filtro che scorre separatamente su ogni mappa, riducendo tutte le quantità che si trovano in quella sezione in un singolo valore. La risoluzione dell'input viene diminuita in base alle dimensioni e allo *stride* del filtro di *pooling*. Le tre tecniche più comuni sono le seguenti:

  - *Max pooling*: viene selezionato il valore massimo da ogni sezione della *feature map* e passato al prossimo step. È il metodo più usato.
  - *Average pooling*: in questo caso viene passato al prossimo livello la media dei valori della sezione.
  - *Global pooling*: viene utilizzato per ricondurre l'intera mappa di una caratteristica a un singolo valore (massimo o medio) prima dello strato finale.

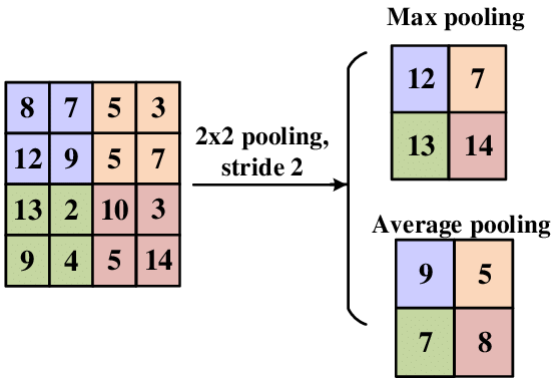

Figura 6: Esempi di risultati dalle operazioni di *max pooling* e *average pooling*.

- Strati completamente connessi: prima di arrivare a questo ultimo step delle CNN, ogni *feature map* proveniente dai livelli precedenti (di *pooling* o convoluzione) viene "appiattita" in un singolo vettore, in modo che possa essere usato come input dei livelli completamente connessi, dove avviene la classificazione. I nodi di questi strati sono connessi a tutti quelli

del livello precedente, sono così in grado di integrare le proprietà estratte ed elaborare le informazioni in entrata in maniera globale [11].

Per concludere, il vettore in uscita da questo strato, passa per una funzione di attivazione (la più comune è *softmax*) nel livello di output, in cui ogni nodo rappresenta una classe, che può essere attivata, se è associata alla probabilità massima [16].

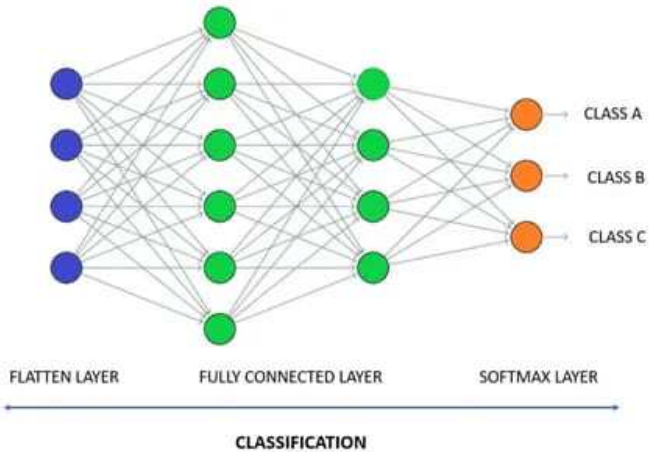

Figura 7: Livelli finali di classificazione. *Feature map* appiattita come input dei livelli completamente connessi e funzione di attivazione *softmax* per l'associazione a classi diverse.

È così spiegata la sequenza di elaborazione delle informazioni, attraverso l'efficiente architettura delle reti neurali convoluzionali, sempre più utilizzate in patologia digitale.

Tornando all'ambito medico, l'immagine istologica digitale originale può quindi essere processata ed etichettata secondo la classe più probabile, grazie ai modelli di CNN, che sono in grado di fare una diagnosi accurata.

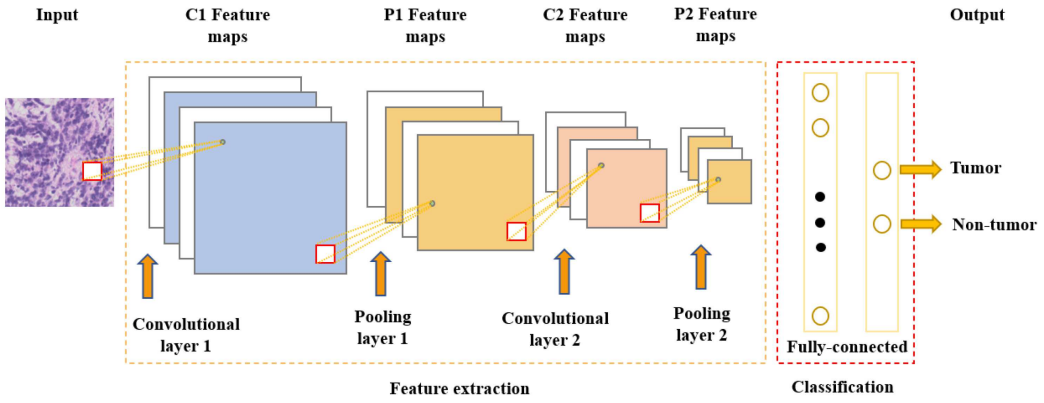

Figura 8: Architettura di una CNN. L'immagine istologica digitale passa attraverso strati di convoluzione, di *pooling* e livelli completamente connessi per classificarla in tessuto tumorale o meno.

Capitolo 3
Classificazione automatica di immagini istopatologiche

La classificazione automatica delle immagini istologiche nella patologia digitale, è una delle applicazioni dei sistemi di deep learning più promettenti. In questo capitolo, nel primo paragrafo, viene descritta l'importanza della disponibilità di dataset, la loro preparazione e annotazione. In seguito, sono esposte le fasi fondamentali del processo di sviluppo di un modello di intelligenza artificiale per l'analisi di immagini biomedicali. Si inizia con la fase di training, in cui vengono evidenziate le principali strategie e tecniche di apprendimento, oltre ai vari metodi di classificazione. La seconda fase è quella di validazione, basata su metriche di valutazione delle prestazioni ben definite e per concludere la fase finale di test dei sistemi automatici in ambito di diagnosi istopatologica.

3.1 Etichettatura e dataset

Come anticipato nei capitoli precedenti, l'addestramento dei sistemi di deep learning richiede un grande numero di dati. Le reti neurali artificiali devono essere in grado di riconoscere efficacemente i dettagli rilevanti da immagini digitali diverse per la diagnosi di patologie, necessitano perciò di ampi dataset di immagini da cui studiare le caratteristiche.

Per un approccio di apprendimento supervisionato (esposto meglio nel prossimo paragrafo), in cui le immagini vengono classificate in base a dei dati *ground-truth*, cioè verificati da patologi, sono necessari dei dataset etichettati.

Campioni di tessuti vengono prelevati tramite biopsia e colorati con la tecnica *H&E* (utilizza due coloranti, Ematossilina e Eosina), che ne permette la visualizzazione ad alta risoluzione al microscopio e la scansione WSI in immagini istologiche digitali [10]. A questo punto, avviene la laboriosa fase di etichettatura: gli esperti patologi assegnano ad ogni immagine per intero, o a delle sezioni (*patch*) o singoli elementi del campione, una o più classi di riferimento; queste saranno prese come verità empiriche. I patologi identificano le regioni di interesse (ROI) e le classificano in base alle loro conoscenze istologiche, alla presenza di lesioni e al grado di severità di un carcinoma. Questa metodologia è sicuramente influenzata da una certa variabilità inter-osservatore.

I dataset possono essere privati o pubblici, tra quelli più rilevanti e che hanno avuto un impatto importante nello sviluppo dei modelli di deep learning, troviamo:

- *CAMELYON* (*Cancer Metastases in Lymph Nodes*): serie di competizioni computazionali internazionali sulla diagnosi automatica da immagini WSI, di metastasi di carcinoma mammario nei linfonodi. Queste challenge si basano su dataset di importanti dimensioni, tra cui *CAMELYON16*: questo archivio fornisce migliaia di immagini istologiche dettagliate e annotate. In questa competizione gli algoritmi di deep learning si sono dimostrati superiori rispetto la diagnosi da parte di esperti patologi [17].
- *PANDA* (*Prostate Cancer Grade Assessment*): quest'altro database, utilizzato per l'addestramento di modelli DL nell'omonima competizione, è uno dei più grandi e si basa su più di 10.000 immagini WSI annotate di biopsie. È focalizzato sulla classificazione, secondo punteggio di Gleason, del cancro alla prostata [6].

I modelli devono essere addestrati su dataset estesi, bilanciati e ben diversificati per tipi di immagini, in modo da rappresentare la grande varietà di possibili casi clinici e rendere il sistema di diagnosi più robusto, oltre che generalizzabile.

Questi dataset sono importanti per l'addestramento dei sistemi DL, che svolgono diversi compiti (come la classificazione, il rilevamento di lesioni o la segmentazione); ma sono anche utili come standard condivisi per la valutazione e il confronto delle prestazioni di architetture diverse.

3.2 Training e modalità di apprendimento

Nella prima fase di training, attraverso un processo iterativo di *backpropagation*, il modello impara a riconoscere pattern, elementi rilevanti e relazioni tra immagini di input e le classi di output, aggiornando continuamente i parametri (pesi) del modello stesso per minimizzare l'errore di predizione. Questa differenza tra valore previsto e valore reale viene calcolata dal modello tramite una funzione di perdita (*loss function*) per ottimizzare il processo.

Gli algoritmi di deep learning possono essere addestrati con modalità diverse.

Possiamo distinguerle in base al tipo di dati utilizzati (metodo supervisionato, non-supervisionato o semi-supervisionato) o da un punto di vista operativo concreto (*from scratch*, *transfer learning* e *fine-tuning*).

3.2.1 Apprendimento supervisionato e non-supervisionato

Nell'apprendimento supervisionato, il modello viene addestrato con dataset etichettati dai patologi e dovrà predire la classe a cui, più probabilmente, appartiene l'input.

Al contrario, nell'apprendimento non-supervisionato, il modello non ha etichette corrette predefinite su cui basarsi, ma deve trovare pattern e correlazioni negli input per conto proprio,

in modo da poter classificare le immagini. Modalità intermedie, come l'apprendimento semi-supervisionato, utilizzano un piccolo dataset etichettato, oltre che a un gran numero di dati non etichettati; questo combina i vantaggi di entrambi i metodi (precisione e scoperta di nuovi dettagli non catturati dall'uomo), migliorandone le prestazioni [14].

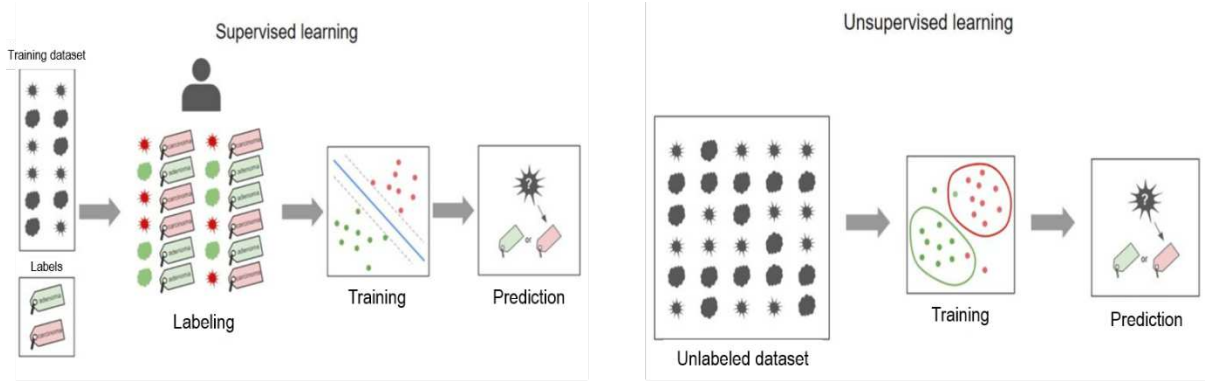

Figura 9: Modalità di apprendimento supervisionato (a sinistra) e non-supervisionato (a destra).

Un interessante strumento di training, sono le *Generative Adversarial Networks* (GAN). Queste particolari reti neurali servono a creare nuovi dati artificiali simili a quelli del dataset originale. Questo sistema è composto da due reti neurali convoluzionali in competizione: una è il generatore, che crea nuovi campioni artificiali, l'altra è il discriminatore, che simultaneamente distingue i dati reali da quelli appena generati. In questo esempio di addestramento non-supervisionato, le due reti competono tra loro, migliorandosi a vicenda, generando immagini sempre più realistiche e imparando sempre più dettagli rilevanti per la discriminazione, minimizzando l'errore di classificazione. Questa tecnica viene anche utilizzata per aumentare la quantità di immagini campione e ampliare il dataset (*data augmentation*) [5, 14].

3.2.2 Learning from scratch e transfer learning

Da un punto di vista più concreto, due sono le principali strategie con cui i modelli di deep learning possono essere addestrati.

Apprendimento da zero (*from scratch*): il sistema viene addestrato dall'inizio, appunto da zero. I parametri, inizializzati casualmente, vengono poi aggiustati a mano a mano che la rete impara i dati appresi nei vari livelli. Questo procedimento possiede sia vantaggi che limiti: è più accurato, ma richiede un ampio dataset iniziale e un grande sforzo computazionale.

Transfer learning: letteralmente apprendimento per trasferimento, infatti, utilizza un modello precedentemente addestrato su un dataset molto ampio, come base per il nuovo compito. Attraverso la strategia di *fine-tuning*, cioè la messa a punto, i valori dei pesi (principalmente

negli strati finali) del modello base vengono modificati e adattati alle nuove richieste più complesse. Nonostante possa risultare meno accurato, questo metodo presenta vari vantaggi e sembra promettente in ambito patologico, in quanto richiede meno tempo e meno dati iniziali (utile poiché i dataset di WSI possono essere limitati e complessi) [15].

3.3 Metodi di classificazione

Riguardo la parte di classificazione di immagini istopatologiche, essa può essere svolta in diverse modalità, in base all'obiettivo diagnostico.

Innanzitutto, la classificazione può essere binaria, per la distinzione in due categorie (ad esempio tessuto sano e tumorale), multiclasse, se il modello assegna all'immagine un'etichetta su tante classi disponibili (ad esempio per diversi tipi o gradi di tumore) o multilabel, se un campione può appartenere contemporaneamente a più classi (ad esempio nel caso di lesioni).

Più nello specifico, la classificazione di immagini istologiche può essere condotta con strategie spaziali diverse: è detta *tile-level,* se avviene su singole porzioni (*patch*), *slide-level*, se viene data un'etichetta globale all'intero vetrino digitale. Possono anche essere usati i cosiddetti approcci *Multiple Instance Learning* (MIL) se all'immagine WSI (di grandi dimensioni) viene assegnata una classe, a partire dall'aggregazione delle informazioni dei singoli *tiles* (porzioni di dimensioni inferiori).

In alcuni casi, la classificazione può inoltre essere preceduta o integrata da un processo di segmentazione delle strutture di interesse, per fornirne la localizzazione precisa nell'immagine. Un esempio concreto è il sistema *U-Net*, un modello specializzato che consente la mappatura dell'area tumorale nel tessuto.

Poiché i dataset istopatologici sono spesso sbilanciati, a causa di classi con quantità diverse di campioni disponibili, vengono utilizzate alcune tecniche per ottimizzare la classificazione e migliorare il riconoscimento di casi rari: *oversampling*, sovracampionamento delle classi minoritarie e *undersampling*, sottocampionamento delle classi maggioritarie, per equilibrare il dataset, oppure l'utilizzo di tecniche di *data augmentation*.

Il processo di classificazione può essere ulteriormente migliorato con la calibrazione delle soglie decisionali e delle probabilità oppure con la combinazione di più modelli [18].

3.4 Validazione e metriche di valutazione

La fase di validazione avviene durante l'addestramento, per monitorare le prestazioni in tempo reale, in modo da individuare subito eventuali problemi ed aggiustare il modello. Usando alcune metriche di valutazione specifiche, è possibile decidere quando fermare l'addestramento e come

ottimizzare i parametri, in modo da ottenere il miglior compromesso tra accuratezza e robustezza. Questa fase di valutazione permette di decidere quali metodi non siano efficienti e debbano essere scartati e quali modelli promettenti possano invece essere migliorati ulteriormente [18].
È importante condurre la fase di validazione seguendo protocolli diversi da una moltitudine di centri istituzionali (diversi metodi di preparazione dei campioni, scanner, digitalizzazione, eccetera), prima che questi modelli di deep learning vengano testati su svariati dataset e utilizzati nella pratica reale della diagnosi [19].

Come anticipato, vengono utilizzate diverse metriche di valutazione quantitative per giudicare le prestazioni di un modello di machine learning.
Per poter definire questi valori, vengono introdotti innanzitutto alcuni concetti relativi ai risultati della più semplice classificazione di tipo binario (che possono poi essere generalizzati ai casi multiclasse e multilabel). I possibili casi risultanti in seguito alla classificazione binaria sono 4, confrontando l'esito previsto con la verità, in generale: un vero positivo (TP), un falso positivo (FP), un vero negativo (TN) e un falso negativo (FN). I veri positivi/negativi sono i valori il cui risultato predetto è corretto, mentre i falsi positivi/negativi derivano dalla previsione di un esito sbagliato, rispetto alla verità, da parte del modello. Per esempio, nel caso di un modello per la diagnosi di cancro, un vero positivo è un paziente che ha realmente il cancro e viene identificato correttamente, mentre un vero negativo è un paziente correttamente riconosciuto come sano. Un falso positivo è un paziente sano a cui viene erroneamente diagnosticato il cancro e un falso negativo è un paziente affetto da carcinoma, che però non viene rilevato dal modello; è quindi evidente l'importanza di valutare un sistema che sarà applicato per la reale diagnosi di patologie, con specifiche metriche quantitative.
Questi esiti possono essere riassunti visivamente in una tabella chiamata *confusion matrix* (matrice di confusione), in cui è più semplice confrontare le previsioni con i reali valori; nel caso della classificazione binaria questa tabella è una matrice 2x2 [18].

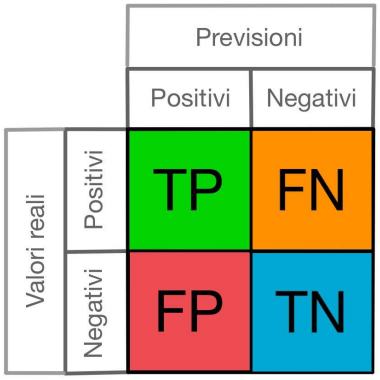

Figura 10: *Confusion matrix*, in cui sono rappresentati i valori di veri positivi (TP), veri negativi (TN), falsi positivi (FP) e falsi negativi (FN).

Tali parametri vengono utilizzati per calcolare le principali metriche di valutazione dei modelli: le loro definizioni e formule sono esposte di seguito.

- Accuratezza: indica la proporzione di predizioni corrette sul totale. È un risultato ingannevole, se il dataset è sbilanciato.

$$Accuratezza = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

- Sensibilità (*Recall*): è la capacità di identificare correttamente i casi positivi. È anche chiamato "tasso di veri positivi" (*true positive rate*).

$$Sensibilità = \frac{TP}{TP + FN} \tag{6}$$

- Specificità: è la capacità di identificare correttamente i casi negativi. È anche detto "tasso di veri negativi" (*true negative rate*).

$$Specificità = \frac{TN}{TN + FP} \tag{7}$$

- Precisione: indica la proporzione delle predizioni positive corrette sul totale dei positivi previsti; cioè tra tutti i campioni classificati come positivi, quanti lo sono davvero. Misura perciò la capacità del modello di evitare falsi positivi. È anche detto "valore predittivo positivo" (*positive predictive value*).

$$Precisione = \frac{TP}{TP + FP} \tag{8}$$

Esistono molti altri indicatori di prestazione, ad esempio:

- *F1-score*: unisce precisione e sensibilità in un unico indice, utile per trovare un equilibrio tra i due valori. Dipende da tutti i parametri della matrice di confusione.

$$F1 = \frac{2 \cdot Precisione \cdot Sensibilità}{Precisione + Sensibilità} \tag{9}$$

- ROC e AUC. La curva ROC si ottiene tracciando il grafico della sensibilità (sull'asse y) in relazione al tasso di falsi positivi (sull'asse x) per diverse soglie di classificazione (la soglia è il valore della probabilità usato come riferimento dal modello, per decidere a quale classe appartenga un dato). Più la curva è vicina all'angolo in alto a sinistra (posizione (0,1)), migliori sono le capacità predittive del modello.

  L'AUC, ovvero la misura dell'area sotto la curva ROC, indica la capacità discriminativa complessiva del modello, a prescindere dalla soglia. Il valore rientra in un range [0,1]; un modello con AUC vicino a 1, ha un'elevata capacità di classificazione [18].

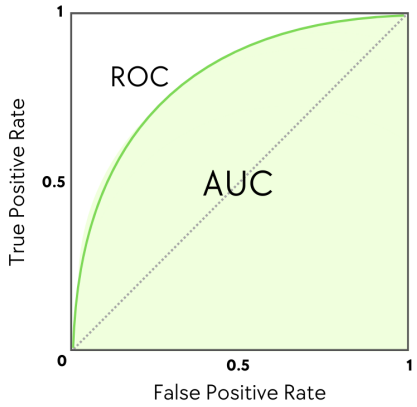

Figura 11: Grafico curva ROC e valore AUC (area sotto la curva).

Le metriche di valutazione sono essenziali per misurare le prestazioni di modelli di apprendimento automatici, per dare informazione sulla loro accuratezza, affidabilità e capacità di generalizzazione. Grazie all'ampio numero di criteri disponibili, è possibile scegliere quelle più adatte in base all'obiettivo dell'analisi, per avere una valutazione completa del sistema.
In ambito medico e patologico, la sensibilità e l'indice F1, sono spesso più importanti dell'accuratezza, perché è fondamentale diminuire il più possibile la diagnosi di pazienti falsi negativi.

3.5 Test

Nella fase finale del processo di sviluppo di un modello di deep learning, dopo che esso è stato addestrato e validato, viene testato su dei dati nuovi, mai utilizzati nelle fasi precedenti. L'obiettivo è quello di valutare realmente le capacità predittive e di classificazione del sistema finale, quindi di scoprire se il modello è generalizzabile su nuovi casi e abbastanza robusto. I dati di test provengono solitamente dallo stesso contesto di raccolta del dataset di training, ma restano separati per garantire un test oggettivo ed imparziale. In questa ultima fase, vengono utilizzate nuovamente le metriche di valutazione standard, per misurare, in questo caso, le

prestazioni finali, verificare la capacità della rete neurale artificiale di generalizzazione su campioni nuovi e valutare come questo si paragoni ad altri modelli e ad esperti patologi [18]. Questo test ci dice quanto il modello sia efficace solo in condizioni sperimentali controllate, non è ancora sufficiente a garantire che il sistema sia applicabile nella pratica clinica quotidiana; saranno necessari ulteriori studi di validazione clinica, che utilizzano dataset esterni reali, di immagini istopatologiche ottenute da laboratori, ospedali e protocolli diversi.

Capitolo 4

Vantaggi, limiti e futuro dei sistemi di deep learning nella diagnostica

L'uso dei sistemi di deep learnig nella diagnostica di immagini istopatologiche digitali è una delle applicazioni più promettenti dell'intelligenza artificiale. Questi modelli automatici permettono l'analisi veloce di grandi quantità di immagini, supportando i patologi esperti nella diagnosi e migliorandone l'accuratezza. D'altro canto, persistono ancora alcuni problemi di generalizzazione, robustezza e soprattutto interpretabilità.

L'obiettivo di questo capitolo è quello di esporre sia i vantaggi dell'utilizzo dei sistemi di deep learning in ambito clinico, che i limiti attuali da superare, evidenziando in particolare alcune tecniche per l'interpretazione dei risultati. Vengono inoltre discusse la necessità di validazione clinica, gli aspetti etici e regolatori e per finire le prospettive future per l'integrazione di questi utili sistemi di diagnosi automatica nella normale pratica clinica.

4.1 Vantaggi

La diagnosi con immagini WSI ad alta risoluzione e modelli di intelligenza artificiale risulta essere utile e promettente in ambito patologico, se non migliore della semplice analisi di tessuti con microscopio, offrendo numerosi vantaggi.

- Formazione ed educazione: le immagini digitali e i sistemi DL sono un utile strumento di insegnamento, offrono supporto alle nuove generazioni di medici, patologi e ricercatori, con annotazioni automatiche e funzioni interattive per un apprendimento dinamico ed efficace.
- Collaborazione e condivisione: questi nuovi strumenti di patologia digitale permettono una più facile collaborazione tra istituti, a prescindere da barriere geografiche, riducendo i tempi di condivisione. Favorisce lo scambio di consulti e seconde opinioni tra esperti a distanza. La condivisione di dataset eterogenei, ottenuti da varie istituzioni, casi e con diversi protocolli, permette inoltre lo sviluppo di modelli di deep learning più robusti e generalizzabili.
- Miglioramenti specifici nella diagnostica: queste tecnologie consentono di ottenere diagnosi con maggiore precisione e riproducibilità inter-osservatore, velocizzano i tempi e riducono i possibili errori umani dal processo. Gli aspetti morfologici dei tessuti vengono misurati più oggettivamente, i modelli AI sono inoltre in grado di riconoscere pattern e strutture più complesse nelle immagini in analisi.

- Informazioni cliniche aggiuntive: i modelli di deep learning possono fornire varie informazioni clinicamente rilevanti, come la predizione dell'esito clinico, la probabilità di recidiva o metastasi di un tumore, oltre alla previsione della risposta terapeutica. Recenti studi hanno dimostrato la capacità di alcuni strumenti DL di derivare informazioni sul profilo genetico del tumore, a partire dall'analisi della sua morfologia, per comprendere meglio la biologia alla base dello specifico carcinoma.
- Assistenza ai patologi: poiché questi modelli sono in grado di elaborare i casi più semplici e di routine velocemente e abbastanza affidabilmente, i patologi possono concentrarsi sui casi più complessi e rari, facendosi assistere ulteriormente dall'intelligenza artificiale. Gli strumenti di deep learning funzionano come sistema di supporto decisionale, migliorando l'accuratezza e la coerenza della diagnosi finale.
- Medicina personalizzata: questi strumenti sono fondamentali per lo sviluppo di pratiche mediche personalizzate e di precisione, in quanto la scelta delle terapie viene adattata alle caratteristiche biologiche e cliniche di ogni singolo paziente.
- Vantaggi economici: la diffusione di queste tecnologie favorisce la sostenibilità dei costi e l'efficienza dei servizi diagnostici nel medio-lungo termine. I benefici economici derivano principalmente dalla riduzione dei tempi di analisi e dalla conseguente ottimizzazione delle risorse umane ed economiche. Inoltre, l'aumento dell'accuratezza diagnostica diminuisce il numero di errori, esami ripetuti e trattamenti non necessari, spesso molto costosi. Tutto questo ha complessivamente un impatto positivo sulle spese del sistema sanitario.

Per concludere, gli strumenti di deep learning contribuiscono positivamente alla standardizzazione dei processi diagnostici, offrendo una migliore prognosi, grazie a predizioni più affidabili e accurate [3, 6].

4.2 Limiti e sfide attuali

Nonostante i vantaggi e le potenzialità degli strumenti di intelligenza artificiale per la diagnosi, la loro applicazione in ambito clinico è limitata da alcuni problemi ancora aperti che gli esperti cercano attualmente di risolvere.

- Costi e necessità pratiche: l'utilizzo di sistemi AI richiede ingenti fondi economici per l'acquisto di strumenti necessari, come scanner e hardware specializzati e potenti e non solo. Servono investimenti per: infrastrutture adatte, l'addestramento del personale, il supporto tecnico, la manutenzione e le licenze di software. Ulteriori costi derivano dalla necessità di un ampio spazio di memoria per le WSI (che sono 1000 volte più grandi di un file ottenuto da raggi X) e di una banda larga per la trasmissione di immagini e connessione

continua con il cloud per l'archiviazione remota. La compressione di immagini, può aiutare a ridurre i costi, ma ne riduce la qualità.

- Dataset e problemi metodologici: una delle principali sfide è ottenere dataset ampi, ben curati ed annotati per l'addestramento dei modelli che richiedono una grande quantità di immagini campione di qualità.

  La presenza di squilibri tra le classi, può provocare importanti problemi di bias che riducono l'accuratezza, la generalizzabilità e l'equità del sistema DL. Questi limiti possono essere causati in particolare dall'utilizzo di dati provenienti da un solo ospedale o con un solo tipo di scanner, oppure da campioni sbilanciati per età, genere, etnia o patologia (può essere difficile trovare immagini di casi anormali).

  Un altro ostacolo può essere la mancanza di verità oggettive e set di validazione standardizzati. La differenza tra la risoluzione con cui è stato sviluppato il sistema di elaborazione AI e risoluzione dello scanner, può causare una riduzione di fedeltà e generalizzabilità degli algoritmi.

  Ulteriore sfida attuale, è quella di creare modelli capaci di elaborare dati complessi e multidimensionali, per avere una visione globale del tumore.

  Questi dati dovrebbero, infine, essere facilmente accessibili, riutilizzabili e non dipendenti da un operatore.
- Un'ulteriore sfida riguarda l'interpretabilità dei modelli automatici, che verrà trattata nel dettaglio nel paragrafo 4.3, includendo anche alcune soluzioni.
- Limiti pratici e organizzativi: solo il processo iniziale di scansione e preparazione delle slide richiede molto tempo e lavoro. Laboratori e personale, spesso, non sono pronti a gestire e manovrare il gran numero di dati immagine necessari. Inoltre in medicina, si usano frequentemente informazioni cliniche aggiuntive a quelle fornite dalle immagini istologiche, come la storia medica, l'età e fattori demografici del paziente, per una corretta diagnosi della patologia.
- Restano anche limiti legati a questioni etiche e regolatorie, trattate nel paragrafo 4.5.

Tutti questi fattori, motivano la relativa lentezza nello sviluppo e nell'applicazione dell'intelligenza artificiale in ambito biomedico e nella patologia digitale. Si tratta infatti di un processo complesso che richiede investimenti, soluzioni sempre più avanzate, interpretabili e generalizzabili, al fronte dello scetticismo e della resistenza all'automatizzazione, da parte dei patologi [4, 5, 6, 14].

4.3 Interpretazione dei risultati

Come anticipato nel paragrafo precedente, uno dei maggiori ostacoli all'adozione in ambito clinico dei sistemi di intelligenza artificiale è dato dal problema dell'interpretabilità. Infatti i modelli di deep learning, in particolare le CNN, sono detti *black-box* in quanto non forniscono una spiegazione chiara del processo decisionale seguito per arrivare ad una classificazione o una previsione. I modelli più complessi (come le reti neurali profonde), assicurano una maggiore accuratezza, ma allo stesso tempo sono meno interpretabili; bisogna quindi cercare un compromesso accettabile tra prestazione e chiarezza decisionale del modello [6].

La mancanza di interpretabilità dei modelli di DL crea scetticismo e sfiducia da parte dei medici patologi, che richiedono una giustificazione chiara per la diagnosi eseguita, in modo da garantire la sicurezza del paziente.

I metodi *handcrafted* di machine learning, cioè basati su caratteristiche predefinite, sono più trasparenti nel processo decisionale, a discapito però dell'accuratezza, in quanto sono sistemi più semplici. Alcuni ricercatori hanno provato a integrare metodi *handcrafted* a degli approcci deep learning: questa fusione rende il processo più comprensibile e interpretabile, ma anche più complesso e lento, oltre che a necessitare di un gran contributo da parte di esperti patologi per la fase iniziale di definizione delle caratteristiche [5].

Sono stati sviluppati approcci di *Explainable AI* (XAI), per rendere più comprensibili i modelli automatici e capire meglio come funzionino i livelli intermedi nascosti. Alcuni studi hanno introdotto metodi post-hoc, che spiegano un modello complesso dopo l'addestramento (senza modificarlo), evidenziando ad esempio tramite mappe di attenzione visive, le aree dell'immagine rilevanti per la decisione presa dalla rete neurale.

Le principali categorie di metodologie per l'interpretabilità, usate per favorire l'applicazione dei sistemi DL nella pratica clinica, sono esposte di seguito.

- Metodi basati sulla retropropagazione (BBM, *Backpropagation-Based Methods*): queste strategie determinano l'importanza di ciascun pixel, per la previsione dell'output, attraverso la retropropagazione delle variazioni di errore, a partire dall'immagine in input. Sono metodi abbastanza veloci, ma le mappe di importanza prodotte sono di qualità inferiore e non determinano la posizione esatta degli elementi rilevanti. Un esempio di BBM sono le *saliency map*, restituiscono appunto delle mappe di importanza, da sovrapporre all'immagine originale per comprendere quali aree hanno contribuito maggiormente alla classificazione. La *Guided Backpropagation*, è una sorta di deconvoluzione, usata per migliorare la visualizzazione delle caratteristiche e quindi la leggibilità delle mappe [20].

- Metodi basati sulla perturbazione (PBM, *Perturbation-Based Methods*): analizzano la variazione dell'output del modello, quando è applicata una perturbazione sull'immagine in entrata; in questo modo si può studiare l'importanza delle diverse aree e caratteristiche sulla classificazione. Le perturbazioni consistono nel modificare o oscurare una parte dell'input. Queste metodologie richiedono però molto tempo. Alcuni esempi sono *LIME* e *SHAP*, che permettono di capire quanto l'assenza o il cambiamento di una regione influenzi la predizione; si concentrano sul contributo delle singole caratteristiche. Spiegano le predizioni, semplificando il modello, grazie ad un'approssimazione localizzata del modello originale più complesso [20].
- Metodi basati sull'attivazione (ABM, *Activation-Based Methods*): queste strategie analizzano l'attivazione dei livelli interni della rete neurale. Un esempio è la tecnica detta *Class Activation Mapping* (CAM) che genera una mappa di attivazione visiva, chiamata *heatmap*. Queste si formano prendendo le mappe di attivazione di un livello convoluzionale profondo e combinandole tramite somma pesata. Si ottiene quindi una rappresentazione visiva a colori delle zone ritenute più importanti dal modello, da sovrapporre all'immagine originale. Le *heatmap* ottenute tramite CAM sono utili in istopatologia per il loro riscontro immediato e interpretabile, hanno però risoluzione limitata e non sono applicabili a tutte le CNN. Altre note tecniche ABM sono le seguenti: *Grad-CAM*, *Smooth Grad-CAM++* e *Feature CAM*. [20]
- Metodi di interpretabilità a livello di pixel (PIL, *Pixel-Level Interpretability*): forniscono una rappresentazione visiva del processo decisionale, molto precisa, a livello di pixel o di piccola regione dell'immagine. Offrono un livello di dettaglio massimo: sono utili per localizzare con precisione le caratteristiche diagnostiche e per evidenziare pattern e strutture microscopiche rilevanti alla previsione. Migliorano l'accuratezza e l'interpretabilità della diagnosi, grazie alla loro precisione, ma possono essere pesanti da un punto di vista computazionale. Un esempio sono le *heatmap* ad alta risoluzione [9].

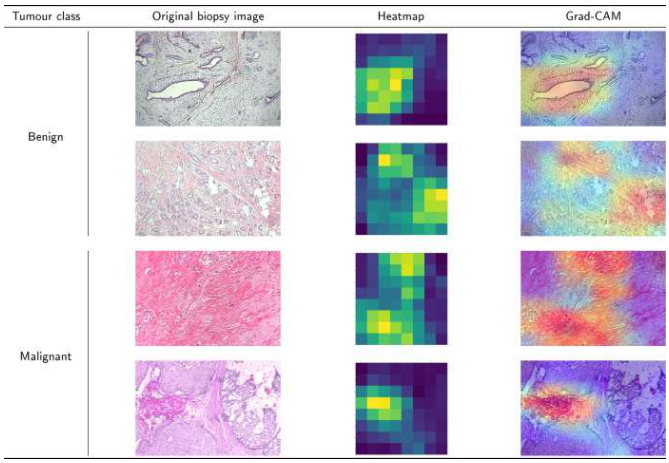

Figura 12: Confronto tra metodi di interpretabilità con *heatmap* (in mezzo) e con *Grad-CAM* (a destra) su immagini di biopsie di tumori benigni (i due campioni superiori) e maligni (i due campioni inferiori).

Si potrebbe però criticare il fatto che un modello AI automatico già complesso di suo, non dovrebbe necessitare di ulteriori sistemi che ne spieghino il funzionamento.

Questi strumenti, in particolare *saliency map* e *heatmap*, permettono di chiarificare e confrontare il ragionamento della rete neurale artificiale con quello del patologo, evidenziando le caratteristiche più importanti per la classificazione e aumentando la fiducia dei medici verso i modelli automatici basati su AI.

In futuro bisognerà creare modelli intrinsecamente più trasparenti, tenendo sempre conto dell'equilibrio da mantenere tra complessità, accuratezza e interpretabilità: i sistemi AI per la diagnosi, verranno accettati più facilmente nella pratica clinica, non solo se avranno ottime prestazioni, ma anche giustificazioni comprensibili e rilevanti per le decisioni prese.

4.4 Validazione clinica

La validazione clinica è un passaggio fondamentale per l'adozione di modelli di deep learning nella pratica diagnostica quotidiana. La validazione clinica si distingue dalla fase di test finale vista precedentemente nell'addestramento in un ambiente sperimentale e controllato, in quanto nella validazione del modello vengono utilizzati dati reali, provenienti da ospedali e pazienti diversi, con strumentazioni e protocolli eterogenei. Questa fase richiede lunghi studi clinici e trial regolatori del modello, oltre alla revisione da parte di esperti patologi, che rappresentano al giorno d'oggi lo standard diagnostico. Esistono comunque alcuni studi che cercano di creare sistemi di standardizzazione più specifici, per un controllo comprensivo e oggettivo della qualità dei sistemi AI diagnostici.

L'obiettivo principale della validazione clinica è quello di verificare robustezza e generalizzabilità del sistema, per assicurare una buona prestazione al di fuori della ricerca sperimentale in laboratorio. È una fase indispensabile a garantire la sicurezza e l'affidabilità per i pazienti, confermando la possibile implementazione del modello AI a supporto dei medici nelle reali decisioni diagnostiche e terapeutiche [6, 19].

4.5 Questioni etiche e regolatorie

Sono stati sollevati numerosi dubbi etici riguardo l'adozione di sistemi di intelligenza artificiale in patologia digitale, che devono essere risolti per poter aumentare la fiducia dei medici e facilitarne l'integrazione nella pratica clinica.

Innanzitutto, per l'utilizzo delle immagini istologiche e altri dati clinici nell'addestramento dei modelli, è necessario il consenso informato da parte dei pazienti; queste sono informazioni sanitarie delicate ed è quindi fondamentale poter garantire sicurezza e anonimato dei dati sensibili delle persone. Esistono normative stringenti a favore della privacy dei pazienti.

La fiducia da parte di pazienti e medici dipende in gran parte dalla trasparenza dei modelli. Questa è inoltre necessaria per ottenere l'autorizzazione per l'utilizzo clinico di strumenti medici automatici dalle autorità regolatorie, che richiedono una descrizione chiara del meccanismo di funzionamento del sistema, in particolare degli algoritmi *black-box* di deep learning. Tali standard severi sono imposti da agenzie, quali FDA (*Food and Drug Administration*) negli Stati Uniti e EMA (*European Medicines Agency*) in Europa. La prima, classifica i modelli AI come dispositivi medici di II o III categoria, ovvero ad alto rischio. Tra i primi strumenti di patologia digitale approvati per il commercio dall'FDA ci sono *Philips IntelliSite Pathology Solution* (nel 2017) e *PAIGE.AI* (2019). Di notevole importanza è il recente passo avanti, dal punto di vista normativo, compiuto da EMA all'inizio del 2025: l'ente ha rilasciato la prima *qualification opinion* (cioè un parere scientifico ufficiale che permette di usare una tecnologia in studi clinici regolati, ma non ne approva ancora il commercio) per un modello AI, chiamato *AIM-NASH*, utile per valutare l'istologia del fegato e il grado di disfunzione metabolica in pazienti affetti da MASH (malattia dovuta all'accumulo di grasso nel fegato, che causa danno epatico e infiammazione) [21].

Vanno considerate inoltre questioni di equità e giustizia sanitaria, dovute a possibili bias degli algoritmi, in quanto, se addestrati su dataset limitati e poco rappresentativi, rischiano di dare diagnosi meno accurate per alcuni sottogruppi di pazienti. Queste tecnologie sono in aggiunta molto costose e quindi non accessibili a tutti.

Andrebbe poi chiarito su chi, tra patologo, sviluppatore del software o agenzia regolatoria, ricadrebbe la responsabilità legale in caso di errori di diagnosi compiuti dal modello AI, che possono avere gravi conseguenze in ambito clinico.
La resistenza al cambiamento e lo scetticismo da parte dei patologi sono quindi motivati dai vari dubbi esposti in precedenza, oltre alla percezione di una perdita di controllo diretto sul processo di analisi diagnostica e la paura che la valutazione clinica venga sostituita completamente da sistemi automatizzati, quando in realtà andrebbero intesi come sistemi di supporto decisionale.
Per velocizzare l'integrazione di modelli AI nella pratica clinica, è necessario stabilire standard univoci per confrontare le prestazioni degli algoritmi e norme chiare per la loro approvazione [6, 19].

4.6 Prospettive future

Tenendo in considerazione i vantaggi esposti all'inizio di questo capitolo, le prospettive future della patologia digitale e dei sistemi di diagnostica basati sul deep learning, sono molto promettenti e, risolvendo gradualmente i problemi ancora aperti, arriveranno a rivoluzionare completamente la pratica medica.
La digitalizzazione delle immagini WSI e lo sviluppo di reti neurali profonde sta trasformando il processo diagnostico attuale, il quale è manuale, qualitativo e soggettivo, in un campo medico più oggettivo e quantitativo, per ottenere diagnosi più rapide ed accurate. I patologi esperti dovranno comunque essere parte integrante del processo decisionale, per supervisionare e validare le tecnologie AI.
Dal punto di vista economico, ci si aspetta che il rapporto costi-benefici migliori sempre di più, ottimizzando i flussi di lavoro e l'allocazione delle risorse. In futuro, saranno però necessari studi per valutare l'adozione di sistemi di patologia digitale in contesti con limitate risorse e in piccoli laboratori.
Lo sviluppo di dataset di immagini digitali sempre più ampi e curati (come quelli sviluppati attualmente dalle aziende *PAIGE.AI* e *Inspirata*), porterà al costante miglioramento del processo di addestramento dei modelli AI, con prestazioni superiori, oltre alla standardizzazione dei processi di diagnosi. Lo sviluppo delle reti 5G e di cloud più accessibili permetterà una migliore gestione e condivisione di questi ampi archivi di grandi immagini digitali.

Come anticipato, gli algoritmi dovranno inoltre migliorare dal punto di vista dell'interpretabilità e della trasparenza nelle decisioni prese, in modo da favorirne l'adozione nella pratica clinica.

In futuro, i modelli AI, attualmente limitati a svolgere compiti specifici, dovranno evolversi in reti neurali multimodali, capaci di integrare informazioni diverse da genomica, proteomica, istopatologia e dati clinici personali, per costruire un profilo dettagliato e comprensivo del paziente e proporre soluzioni terapeutiche personalizzate.

Un'altra rivoluzione tecnologica imminente, riguarda il passaggio dagli approcci di elaborazione 2D all'analisi di immagini 3D dell'intero tessuto, che offrono una visione del volume più completa e realistica. Sarà quindi essenziale adattare gli strumenti AI di analisi istopatologica a questo nuovo tipo di immagini, per poter gestire la grande quantità di dati complessi multidimensionali e aiutare efficacemente i medici nella diagnosi.

Studi recenti hanno dimostrato come la collaborazione tra patologo e algoritmi DL produca diagnosi più accurate, rispetto al risultato del singolo medico o modello. È quindi evidente come i sistemi di deep learning saranno, non un sostituto, ma uno strumento sempre più di supporto per i medici, altamente efficace nella diagnosi di immagini digitali istopatologiche [4, 5, 6].

Conclusione

Lo sviluppo e l'applicazione dell'intelligenza artificiale, in particolare delle reti neurali profonde, in patologia digitale, rappresenta un punto di svolta significativo nella medicina moderna. Questa tesi ha messo in evidenza i processi fondamentali, i vantaggi e i limiti legati ai sistemi di deep learning per la diagnostica.

Innanzitutto, è stato spiegato il passaggio dai metodi tradizionali con microscopio alla digitalizzazione delle immagini istologiche, che insieme alla maggiore potenza computazionale disponibile e alla condivisione di dataset WSI di grandi dimensioni, ha favorito l'uso di algoritmi di intelligenza artificiale nell'elaborazione delle immagini istopatologiche.

Passaggio importante nella tesi è stato spiegare nel particolare il funzionamento delle reti neurali artificiali profonde, evidenziando il flusso e l'elaborazione dei dati nei vari livelli delle CNN. Si è sottolineato come questi modelli siano in grado di apprendere automaticamente le caratteristiche più rilevanti nelle immagini fornite come input, superando i modelli basati su caratteristiche *handcrafted*.

Nel contesto della classificazione automatica di immagini istopatologiche, si è passati dallo spiegare l'importanza della creazione di dataset ben curati, ai metodi di training (apprendimento supervisionato e non, oltre a quello da zero e per trasferimento), di validazione e test. Queste fasi, unite all'utilizzo di metriche standard di valutazione, sono essenziali per ottenere modelli di classificazione accurati, robusti e generalizzabili.

Nell'ultimo capitolo, si sono discussi i lati positivi e negativi dell'adozione di questi strumenti nella partica medica. I principali vantaggi sono legati all'accuratezza e rapidità dei risultati diagnostici, oltre all'evoluzione verso una medicina più oggettiva, quantitativa e personalizzata. Gli ostacoli riguardano invece l'accessibilità dei dati, possibili bias, i costi e la scarsa interpretabilità del processo decisionale dei modelli AI. È quindi necessario sottoporre questi sistemi a validazione clinica, affrontare questioni etiche e regolatorie per aumentare la sicurezza e la fiducia di patologi e pazienti nell'utilizzo di sistemi automatici di diagnosi e prognosi.

Infine, la tesi guarda al futuro dei modelli diagnostici AI, che dovranno evolversi in sistemi multimodali, per un'analisi completa e multidimensionale delle informazioni cliniche dei pazienti, superando ostacoli tecnici ed etici in modo da garantire una sicura integrazione nei processi di analisi diagnostica.

In conclusione, l'intelligenza artificiale in patologia digitale, non va considerata come una minaccia alla professione medica, ma come uno strumento di supporto da affiancare ai patologi, che permette loro di concentrarsi su casi più complicati e ottenere analisi più dettagliate.

Le reti neurali artificiali stanno già trasformando il modo di svolgere diagnosi, classificazioni e prognosi di malattie oncologiche e non solo; potranno realizzare il loro massimo potenziale in futuro, superando sfide ancora aperte, tramite la collaborazione di ricercatori, ingegneri, medici ed istituzioni.

I modelli di deep learning, nel campo della diagnostica istopatologica, rappresentano una vera rivoluzione metodologica, che potrà migliorare il sistema sanitario con analisi più accurate, efficienti ed eque, se tali strumenti continueranno ad essere sviluppati ed implementati con responsabilità.

Sitografia

[1] https://www.humanitas.it/enciclopedia/specialisti/anatomopatologo/

[2] https://www.treccani.it/enciclopedia/istopatologia_(Dizionario-di-Medicina)/

[3] Prasad P, Khair AM, Saeed M, Shetty N. *Artificial intelligence in histopathology*. (2024) DOI:10.4103/jpbs.jpbs_727_24 https://journals.lww.com/jpbs/fulltext/2024/16005/artificial_intelligence_in_histopathology.7.aspx

[4] Mubarak M. *Move from traditional histopathology to digital and computational pathology: Are we ready?* (2022) DOI: 10.4103/ijn.IJN_508_20 https://pmc.ncbi.nlm.nih.gov/articles/PMC9775606/pdf/IJN-32-414.pdf

[5] Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. *Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology.* (2019) https://doi.org/10.1038/s41571-019-0252-y

[6] Shafi S, Parwani AV. *Artificial intelligence in diagnostic pathology.* (2023) https://doi.org/10.1186/s13000-023-01375-z

[7] https://www.agendadigitale.eu/sanita/intelligenza-artificiale-e-digital-pathology-cosa-cambia-per-diagnosi-prognosi-e-cura/

[8] https://www.isahit.com/blog/what-is-the-purpose-of-image-preprocessing-in-deep-learning

[9] Ennab, M, Mcheick H. *Advancing AI Interpretability in Medical Imaging: A Comparative Analysis of Pixel-Level Interpretability and Grad-CAM Mode*. 2025. https://doi.org/10.3390/make7010012

[10] Dabeer S, Khan MM, Islam S. *Cancer diagnosis in histopathological image: CNN based approach.* (2019) https://doi.org/10.1016/j.imu.2019.100231

[11] https://towardsdatascience.com/the-math-behind-convolutional-neural-networks-6aed775df076/#:~:text=At%20its%20core%2C%20the%20convolution,the%20kernel%20is%20currently%20positioned

[12] https://www.innovationpost.it/tecnologie/intelligenza-artificiale-deep-learning-e-machine-learning-quali-sono-le-differenze/

[13] https://www.ibm.com/it-it/think/topics/neural-networks

[14] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez CI. *A survey on deep learning in medical image analysis.* (2017)

https://doi.org/10.1016/j.media.2017.07.005

[15] Rodriguez LF, Naldi MC, Mari JF. *Comparing convolutional neural networks and preprocessing techniques for HEp-2 cell classification in immunofluorescence images.* (2019) https://doi.org/10.1016/j.compbiomed.2019.103542

[16] https://medium.com/@beingfarina/breaking-down-the-mathematics-behind-cnn-models-a-comprehensive-guide-1853aa6b011e

[17] Campanella G, Hanna MG, GeneslawL. et al. *Clinical-grade computational pathology using weakly supervised deep learning on whole slide images.* (2019) https://doi.org/10.1038/s41591-019-0508-1

[18] Rainio O, Teuho J & Klén R. *Evaluation metrics and statistical tests for machine learning.* (2024) https://doi.org/10.1038/s41598-024-56706-x

[19] Ahmed AA, Abouzid M, Kaczmarek E. *Deep Learning Approaches in Histopathology.* (2022) https://doi.org/10.3390/cancers14215264

[20] Clement F, Yang J, Cheng I. *Feature CAM: Interpretability in Image Classification with Improved Visual Attention.* (2024) https://doi.org/10.48550/arXiv.2403.05658

[21] https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/artificial-intelligence?utm_source=chatgpt.com#first-qualification-opinion-on-ai-methodology-71958

Elenco delle figure

(1) https://www.researchgate.net/publication/370507856_hist2RNA_An_Efficient_Deep_Learning_Architecture_to_Predict_Gene_Expression_from_Breast_Cancer_Histopathology_Images

(2) Diagramma creato su Word

(3) https://www.linkedin.com/pulse/back-propagation-algorithm-raji-shanmugam-jjcgc

(4) https://www.researchgate.net/publication/340992802_Convolutional_Neural_Network_Based_Inter-Turn_Fault_Diagnosis_in_LSPMSMs

(5) https://courses.cs.washington.edu/courses/cse446/22wi/sections/08/convolutional_networks.html

(6) https://www.researchgate.net/publication/340812216_Deep_Neural_Networks_on_Chip_-_A_Survey

(7) https://www.aiknow.io/machine-learning-e-time-series/?doing_wp_cron=1755963767.4488730430603027343750

(8) https://www.mdpi.com/2079-3197/11/4/81

(9) https://radiologica.net/modelli-di-machine-learning/

(10) Diagramma creato su Goodnotes

(11) https://www.blog.trainindata.com/auc-roc-analysis/

(12) https://www.sciencedirect.com/science/article/abs/pii/S0045790622005997