



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

Dipartimento di Ingegneria dell'Energia Elettrica e dell'Informazione

“Guglielmo Marconi” – DEI

Corso di Laurea in Ingegneria Biomedica

# **Diagnosi precoce dell'Alzheimer mediante analisi di neuroimmagini e Deep Learning**

*Relatore:*

**Chiar.mo Prof.  
Cristiano Cuppini**

*Presentata da:*

**Michael Di Monte**

---

Sessione Giugno 2026  
Anno Accademico 2025/2026



# Abstract

La Malattia di Alzheimer è una delle principali sfide sanitarie del XXI secolo. La diagnosi nella fase prodromica, conosciuta clinicamente come *Mild Cognitive Impairment*, è un obiettivo fondamentale per massimizzare l'efficacia dei trattamenti cosiddetti *disease-modifying*, ma, spesso e volentieri, l'intercettazione di questa fase risulta complessa a causa delle difficoltà nella rilevazione dei sintomi clinici precoci.

Il presente elaborato si pone come obiettivo l'analisi dell'utilizzo dell'Intelligenza Artificiale, in particolare con un focus sul Deep Learning, come supporto per la diagnosi precoce basata su neuroimmagini.

Cominciando dall'analisi dei fondamenti fisici della principale strumentazione biomedica utilizzata in quest'ambito e dalle pipeline di pre-processing necessarie per elaborare e normalizzare i dati acquisiti, questa tesi analizza l'evoluzione dei modelli di classificazione e degli algoritmi utilizzati maggiormente: a partire dagli approcci classici di Machine Learning, fino ad arrivare alle moderne Reti Neurali Convolutionali, o CNN, capaci di estrarre automaticamente e autonomamente le feature comunemente associate alla neurodegenerazione.

A seguito di un'analisi della letteratura scientifica e dello stato dell'arte, vengono discussi i risultati ottenuti da diverse ricerche su dataset standardizzati, come ADNI, e si evidenzia il vantaggio ottenibile da approcci multimodali.

Infine, vengono discussi i principali ostacoli che rendono difficile l'implementazione clinica, come la scarsa interpretabilità dei modelli e le difficoltà di generalizzazione su dati eterogenei, dando uno sguardo sulle prospettive future nello sviluppo di sistemi di Computer-Aided Diagnosis che possano garantire affidabilità e trasparenza decisionale.



# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
1.1	Il contesto clinico: la malattia di Alzheimer . . . . .	1
1.2	La sfida della diagnosi precoce: dallo stato Cognitivo Normale al Disturbo Cognitivo Lieve . . . . .	1
1.3	Il ruolo del neuroimaging come biomarker quantitativo . . . . .	2
1.4	L'avvento dell'Intelligenza Artificiale per l'analisi biomedica . . . . .	3
1.5	Obiettivi e struttura dell'elaborato . . . . .	4
<b>2</b>	<b>Fondamenti di Strumentazione Biomedica per il Neuroimaging dell'AD</b>	<b>5</b>
2.1	La Risonanza Magnetica . . . . .	5
2.1.1	Principi fisici e generazione dell'immagine . . . . .	5
2.1.2	MRI Strutturale per l'analisi morfometrica . . . . .	7
2.1.3	MRI Funzionale e di Diffusione per l'analisi della connettività . . . . .	8
2.2	La Tomografia a Emissione di Positroni . . . . .	8
2.2.1	Principi fisici e traccianti . . . . .	9
2.2.2	FDG-PET per la valutazione dell'ipometabolismo . . . . .	9
2.2.3	Amyloid-PET e Tau-PET per i biomarker specifici . . . . .	10
2.3	La Tomografia Computerizzata e il suo ruolo . . . . .	10
2.3.1	Applicazione clinica nell'Alzheimer . . . . .	11
2.4	Il pre-processing delle immagini: dalla strumentazione al dato analizzabile . . . . .	11
2.4.1	Rimozione degli artefatti e Skull Stripping . . . . .	11
2.4.2	Correzione delle disomogeneità . . . . .	12
2.4.3	Registrazione e Normalizzazione . . . . .	12
2.4.4	Software di riferimento nella ricerca . . . . .	14
<b>3</b>	<b>Metodologie di Intelligenza Artificiale per la Classificazione</b>	<b>15</b>
3.1	Approcci di Machine Learning tradizionale . . . . .	15
3.1.1	Estrazione manuale delle feature . . . . .	16
3.1.2	Support Vector Machines . . . . .	17
3.2	Approcci di Deep Learning . . . . .	18
3.2.1	Le Reti Neurali Convoluzionali per l'analisi di immagini . . . . .	19

3.2.2	Architetture comuni e Transfer Learning . . . . .	21
3.3	Metriche di valutazione dei modelli . . . . .	22
3.3.1	Metriche standard: Accuratezza, Sensitività, Specificità, Precisione, Curva ROC/AUC . . . . .	22
3.3.2	Strategie di validazione robusta: Cross-Validation e prevenzione del Data Leakage . . . . .	23
<b>4</b>	<b>Analisi dello Stato dell'Arte</b>	<b>25</b>
4.1	Rassegna delle applicazioni di AI su dati MRI . . . . .	25
4.2	Rassegna delle applicazioni di AI su dati PET . . . . .	26
4.3	Approcci multimodali: la fusione di dati da strumentazione diversa . . . . .	27
4.4	Dataset di riferimento . . . . .	28
4.5	Le sfide emergenti . . . . .	29
4.5.1	Interpretabilità e Explainable AI . . . . .	29
4.5.2	Generalizzabilità e Domain Shift . . . . .	29
<b>5</b>	<b>Discussione e Conclusioni</b>	<b>31</b>
5.1	Sintesi dei risultati e confronto delle metodologie . . . . .	31
5.2	Criticità e limiti della ricerca attuale . . . . .	32
5.3	Prospettive future: l'integrazione nella pratica clinica . . . . .	33
5.4	Considerazioni finali . . . . .	34
	<b>Bibliografia</b>	<b>35</b>

# Capitolo 1

## Introduzione

### 1.1 Il contesto clinico: la malattia di Alzheimer

La Malattia di Alzheimer (AD) è una patologia neurodegenerativa irreversibile e incurabile, che rappresenta il 60-70% dei casi di demenza, rendendola di fatto la forma più comune [1]. Storicamente, l'Alzheimer veniva definito basandosi sui sintomi clinici classici della demenza; tuttavia, attualmente ci si è evoluti verso un concetto di "continuum" biologico che si estende per la durata di 15-25 anni [2] e che si può ritenere progressivo. Dal punto di vista epidemiologico, l'AD è una delle sfide sanitarie più costose e critiche del secolo attuale, insieme alle altre forme di demenza; a conferma di quanto detto, le proiezioni attuali indicano che il numero di persone affette da demenza raddoppierà in Europa e triplicherà a livello globale [2].

Il quadro clinico classico emerge solitamente dopo i 60 anni di età, nonostante possano comunque presentarsi forme a esordio precoce o varianti atipiche in età più giovane [1]; i sintomi tipici includono deficit nella memoria episodica a breve termine, seguiti da disfunzioni esecutive e visuo-spaziali che portano inevitabilmente alla perdita di indipendenza funzionale [1].

I criteri diagnostici storici prendevano in considerazione i sintomi sopracitati, definendo l'AD sulla base di questi [3]. Tuttavia, attendere il declino cognitivo per formulare una diagnosi rendeva impossibile l'intervento tempestivo, rendendo di fatto necessario esplorare le fasi antecedenti alla demenza conclamata.

### 1.2 La sfida della diagnosi precoce: dallo stato Cognitivo Normale al Disturbo Cognitivo Lieve

Nel modello di continuum del declino cognitivo, il Disturbo Cognitivo Lieve (*Mild Cognitive Impairment* – MCI) viene considerato come lo stadio sintomatico pre-demenza. Esso si posiziona a metà strada tra l'invecchiamento fisiologico e la demenza conclamata e si stima che il

10-20% degli adulti di età superiore ai 65 anni ne sia affetto, con un rischio che aumenta in maniera direttamente proporzionale rispetto all'età [4]. Il MCI è caratterizzato principalmente dalla compromissione di uno o più domini cognitivi, come per esempio memoria, funzioni esecutive, attenzione o linguaggio, ma con un sostanziale mantenimento dell'indipendenza nelle attività di vita quotidiana (ADL - *Activities of Daily Living*) [4].

A causa di tale conservazione dell'autonomia funzionale, la diagnosi precoce in questa fase rappresenta una sfida clinica molto complessa: attualmente, infatti, il MCI presenta alti tassi di sottodiagnosi o di diagnosi errate a causa di limiti operativi come la breve durata delle tipiche visite mediche, la mancanza di formazione specifica dei medici di base e l'assenza di screening rapidi e sensibili [5]. Il problema principale derivante dalla mancata identificazione nella fase precoce limita l'efficacia delle strategie terapeutiche *disease-modifying* e comporta il rischio di precluderne l'accesso, poiché quest'ultime richiedono un intervento nelle fasi prodromiche della malattia [5]; alcuni studi hanno però evidenziato come i pazienti con MCI abbiano profili clinici distinti, caratterizzati da sottili – ma identificabili – deficit nelle ADL, come per esempio difficoltà nella gestione delle finanze o nel ricordare appuntamenti, in aggiunta a una maggiore frequenza di sintomi neuropsichiatrici quali *apatia*, irritabilità o *disforia* [6].

La prognosi del MCI è spesso soggetta a grande incertezza: sebbene coloro che ne soffrono abbiano un rischio di conversione in demenza superiore alla media, la variabilità è elevatissima, con tassi di conversione annuale ampiamente variabili a seconda della popolazione studiata e dai criteri diagnostici applicati [4]; bisogna infatti notare che una percentuale significativa di soggetti potrebbe non progredire verso la demenza anche dopo 10 anni, o addirittura regredire a uno stato Cognitivo Normale [4]. Questa variabilità ha reso necessario un cambio di paradigma diagnostico: dal 2018, infatti, l'AD viene definito dalla presenza di specifiche alterazioni chimiche e fisiche nel corpo e nel cervello; il passaggio dalla definizione clinica a quella biologica è supportata dal framework ATN (Amiloide, Tau, Neurodegenerazione): un sistema standardizzato che utilizza tre categorie di biomarcatori per stabilire la presenza o meno dell'AD tramite PET e analisi del liquido cerebrospinale (CSF) [7].

### **1.3 Il ruolo del neuroimaging come biomarker quantitativo**

A livello fisiopatologico, la malattia di Alzheimer è caratterizzata dall'accumulo extracellulare di placche di proteina  $\beta$ -Amiloide ( $A\beta$ ) e dalla formazione intracellulare di grovigli neurofibrillari di proteina Tau iperfosforilata [1]. Queste alterazioni rappresentano quella che viene chiamata la "fase cellulare" della malattia, che precede i primi sintomi. Il framework NIA-AA del 2018 ha descritto formalmente questa fase con il sistema ATN (Amiloide, Tau, Neurodegenerazione), affermando che la diagnosi vera e propria si ha con la presenza simultanea dei fattori A e T (A+T+); invece, la sola positività all'Amiloide (A+T-) è classificata come "cambiamento

patologico associato all'Alzheimer". È chiaro quindi come la proteina Tau sia il marker critico per definire la malattia conclamata [7].

In questo contesto, il neuroimaging rappresenta una fonte molto importante di biomarker quantitativi in modo da monitorare i processi fisiopatologici *in vivo*. La Risonanza Magnetica Strutturale (sMRI) è quindi fondamentale per determinare la neurodegenerazione, in quanto l'atrofia cerebrale segue pattern specifici che coinvolgono l'ippocampo, il lobo temporale mediale e la corteccia entorinale [8]; allo stesso tempo, la PET, specificamente la variante FDG-PET, consente di mappare il metabolismo cerebrale, permettendo di individuare e mappare i processi disfunzionali con precisione, in quanto le alterazioni neuronali spesso precedono l'atrofia strutturale visibile alla sMRI [8].

Recentemente, i criteri NIA-AA sono stati aggiornati ulteriormente per integrare i biomarcatori ematici, chiamati *Blood Based Biomarkers* o BBM, come la proteina p-tau217, che permettono uno screening non invasivo associato ad un alto rendimento [9]. Tuttavia, sebbene i BBM siano una soluzione rivoluzionaria per lo screening di massa grazie alla loro non invasività e ai costi ridotti, essi forniscono informazioni puramente biochimiche: un test ematico può quindi confermare la positività alle proteine patogene, diventando quindi un indicatore di rischio biologico, ma non può descriverne l'impatto o localizzare i danni. Il neuroimaging, di conseguenza, rimane insostituibile per la conferma della neurodegenerazione e del danno strutturale. Nello specifico, il poter identificare pattern di atrofia nell'ippocampo e nel lobo temporale mediale rimane fondamentale per formulare una diagnosi accurata [9].

La necessità di estrarre queste informazioni, spesso tanto complesse ed eterogenee da essere impercettibili all'occhio umano, soprattutto nella fase prodromica della malattia, richiede lo sviluppo di metodologie di analisi capaci di gestire questi dati in modo da garantire una diagnosi di precisione.

## **1.4 L'avvento dell'Intelligenza Artificiale per l'analisi biomedica**

L'ultimo decennio ha segnato un punto di svolta nella ricerca medica grazie ad un'enorme integrazione dell'Intelligenza Artificiale (AI), campo che ha registrato una crescita esponenziale in termini di pubblicazioni scientifiche e investimenti, soprattutto negli ultimi cinque anni, grazie alla disponibilità di *big data* e all'aumento della potenza di calcolo [10].

L'AI, con le sue sottocategorie come il Machine Learning (ML) e il Deep Learning (DL), permette di simulare processi cognitivi veri e propri, rendendo le macchine capaci di apprendere

dall'esperienza per compiti come diagnosi, prognosi e scoperta di nuovi biomarcatori [10]. Questo ha rivoluzionato la ricerca tradizionale, grazie anche all'abbandono dei sistemi algoritmici classici basati su regole fisse (*rule-based*) in favore di modelli addestrati sui dati (*data-driven*) [11]. Queste nuove tecnologie sono in grado di elaborare dataset complessi con una velocità e una precisione elevatissime, migliorando sia l'accuratezza diagnostica che le tempistiche della ricerca, aspetto cruciale in patologie come l'Alzheimer [12]. Nel contesto specifico dell'AD, applicare l'AI alle neuroimmagini, in particolare MRI, ha portato allo sviluppo di algoritmi sofisticati per la diagnosi precoce e per la predizione dello sviluppo della malattia [13].

Negli anni più recenti, abbiamo assistito a un'evoluzione degli algoritmi di AI: se gli approcci di Machine Learning tradizionale richiedevano un'estrazione manuale delle feature dalle immagini, l'avvento del Deep Learning ha introdotto reti neurali capaci di apprendere automaticamente le feature gerarchiche direttamente dai dati grezzi [13]. Questi modelli permettono, con prestazioni elevate, di classificare tra soggetti sani, MCI e AD, rilevando pattern di atrofia microstrutturale che altrimenti sarebbero impercettibili all'occhio umano, permettendo un intervento tempestivo [13].

## 1.5 Obiettivi e struttura dell'elaborato

L'obiettivo di questo lavoro di tesi è analizzare il ruolo dell'Intelligenza Artificiale, nello specifico del Deep Learning, come strumento di supporto alla diagnosi precoce dell'Alzheimer, valutandone l'efficacia nell'elaborazione di neuroimmagini a tale scopo.

La trattazione si apre nel secondo capitolo, in cui verrà fatta una disamina dei fondamenti fisici della principale strumentazione biomedica utilizzata in questo contesto specifico, ovvero MRI e PET, con uno sguardo sui protocolli di pre-processing necessari per normalizzare i dati per renderli input idonei per gli algoritmi di AI.

A partire da questi fondamenti tecnici, nel terzo capitolo verranno analizzate le architetture di classificazione maggiormente utilizzate, cominciando dal Machine Learning tradizionale, fino ad arrivare alle più recenti e avanzate Reti Neurali Convoluzionali, definendo le metriche matematiche per la valutazione delle performance di questi algoritmi.

Il quadro si completa nel quarto capitolo, in cui viene fatta una rassegna sia dello stato dell'arte in letteratura, che dei risultati di test eseguiti su dataset standardizzati, evidenziando le criticità ingegneristiche ancora irrisolte riguardo l'interpretabilità dei modelli e la loro generalizzabilità clinica.

Infine, il quinto capitolo ha l'obiettivo di sintetizzare i risultati raccolti, analizzando e discutendo le prospettive future per l'integrazione di queste tecnologie nella pratica medica comune.

## Capitolo 2

# Fondamenti di Strumentazione Biomedica per il Neuroimaging dell'AD

Il continuo miglioramento tecnologico nell'ambito della strumentazione biomedica ci permette di esplorare il cervello umano più a fondo e con precisione maggiore. Questo apre nuove strade nella comprensione e nella diagnosi precoce della malattia di Alzheimer, ed è così che le tecniche di neuroimaging diventano strumenti essenziali per mostrare i processi di degenerazione cerebrale nascosti, trasformando dati complessi in informazioni cliniche chiare e utili. Di seguito le tecnologie maggiormente utilizzate per la diagnosi e nel monitoraggio dell'AD e qualche cenno sul loro impiego nell'addestramento di modelli di AI.

### 2.1 La Risonanza Magnetica

La Risonanza Magnetica (MRI – *Magnetic Resonance Imaging*) rappresenta ad oggi il gold standard per l'indagine non invasiva dell'encefalo, grazie alla sua elevata risoluzione spaziale e all'eccellente contrasto che mostrano i tessuti molli. A differenza della Tomografia Computerizzata, non usa radiazioni ionizzanti, ma sfrutta le proprietà magnetiche dei nuclei atomici, nello specifico di quelli dell'idrogeno ( $^1\text{H}$ ), abbondantemente presenti in acqua e nei lipidi nel corpo umano, per generare immagini tomografiche a tre dimensioni [14].

Nel contesto dell'Alzheimer, la MRI aiuta a misurare l'atrofia cerebrale e a quantificare i danni alla connettività neuronale [15], fornendo i dati strutturali e, nel caso della fMRI, funzionali, per addestrare i modelli di Deep Learning [16].

#### 2.1.1 Principi fisici e generazione dell'immagine

Il principio fisico alla base del funzionamento della MRI è l'interazione tra il momento magnetico dei nuclei e un campo magnetico esterno. In condizioni di riposo, i protoni dell'idrogeno presentano un momento magnetico, o spin, orientato casualmente; quando il paziente viene

introdotto nel magnete principale dello scanner, che genera un campo magnetico statico e omogeneo  $B_0$  (tipicamente di 1.5 T o 3 T per uso clinico e fino a 7 T per la ricerca), gli spin nucleari tendono ad allinearsi parallelamente o antiparallelamente alla direzione del campo; questo allineamento genera una magnetizzazione netta  $M$  lungo l'asse longitudinale  $z$ . I nuclei, inoltre, non sono statici ma precessano attorno alla direzione di  $B_0$  con una frequenza angolare ben specifica, definita dall'equazione di Larmor [14]:

$$\omega_0 = \gamma B_0 \quad (2.1)$$

dove  $\omega_0$  è la frequenza di precessione, o frequenza di Larmor, mentre  $\gamma$  è il rapporto giromagnetico, che rappresenta il rapporto tra il momento magnetico e il momento angolare ed è una costante caratteristica intrinseca (per l'idrogeno  $\gamma/2\pi \approx 42.58 \text{ MHz T}^{-1}$ ). La generazione del segnale avviene attraverso tre fasi principali [14]:

1. **Eccitazione:** una bobina a radiofrequenza trasmette un impulso elettromagnetico alla frequenza di Larmor  $\omega_0$  (campo  $B_1$ ). Questo impulso trasferisce energia ai nuclei, ribaltando il vettore della magnetizzazione netta dal piano longitudinale al piano trasversale  $xy$ .
2. **Rilassamento:** al termine dell'impulso, il sistema viene lasciato tornare all'equilibrio energetico. Questo processo è descritto da due costanti di tempo, entrambe caratteristiche del tessuto:
  - Rilassamento Longitudinale ( $T_1$ ): o rilassamento spin-reticolo, è il tempo necessario affinché la componente longitudinale della magnetizzazione ( $M_z$ ) recuperi il 63% del suo valore iniziale. Dipende dalla capacità del nucleo di cedere energia all'ambiente circostante, il "reticolo".
  - Rilassamento Trasversale ( $T_2$ ): o rilassamento spin-spin, è il tempo necessario affinché la componente trasversale della magnetizzazione ( $M_{xy}$ ) decada al 37% del valore massimo, causato dalla perdita della coerenza di fase tra gli spin.
3. **Acquisizione e Codifica Spaziale:** Durante la fase di rilassamento, la bobina capta il segnale generato (o FID – *Free Induction Decay*). Per generare l'immagine, bisogna localizzare il segnale nello spazio: ciò è possibile applicando gradienti lineari di campo magnetico nelle tre dimensioni ( $G_x, G_y, G_z$ ), i quali modificano la frequenza di risonanza in funzione della posizione. I dati vengono quindi acquisiti nel dominio delle frequenze spaziali, chiamato *K-space*, mentre in seguito, tramite una Trasformata di Fourier Veloce (FFT – *Fast Fourier Transform*), vengono trasformati in immagine. Manipolando poi i parametri di acquisizione, in particolare il Tempo di Ripetizione ( $TR$ ) e il Tempo di Eco ( $TE$ ), si possono esaltare le differenze di  $T_1$  e  $T_2$ , ottenendo sequenze pesate (*weighted*) diverse [14].

### 2.1.2 MRI Strutturale per l'analisi morfometrica

La Risonanza Magnetica Strutturale (sMRI) è la modalità principale per l'analisi morfometrica del cervello nell'Alzheimer. Le sequenze più utilizzate in quest'ambito sono le sequenze pesate in T1 (*T1-weighted*), poiché offrono un migliore contrasto tra materia grigia (GM), materia bianca (WM) e liquido cerebrospinale (CSF) [14]. In un cervello affetto da AD, la neurodegenerazione è l'effetto principale dell'atrofia dei tessuti cerebrali, visibile chiaramente in Figura 2.1; analizzare morfometricamente l'encefalo, di conseguenza, permette di identificare biomarcatori quantitativi cruciali [15]:

- **Atrofia dell'Ippocampo:** è uno dei segni più precoci di Alzheimer. L'ippocampo e la corteccia entorinale, infatti, subiscono una riduzione di volume significativa già nelle prime fasi di MCI.
- **Allargamento dei ventricoli:** come conseguenza diretta della perdita di volume parenchimale, i ventricoli si allargano.
- **Assottigliamento corticale:** si ha una riduzione dello spessore della corteccia cerebrale in determinate zone, soprattutto nei lobi temporali e parietali.

Questi dati possono essere elaborati dagli algoritmi di AI attraverso tecniche come la *Voxel-Based Morphometry* (VBM) o l'estrazione di *Regions of Interest* (ROI) per poter classificare il soggetto come sano (CN), MCI o AD [16].

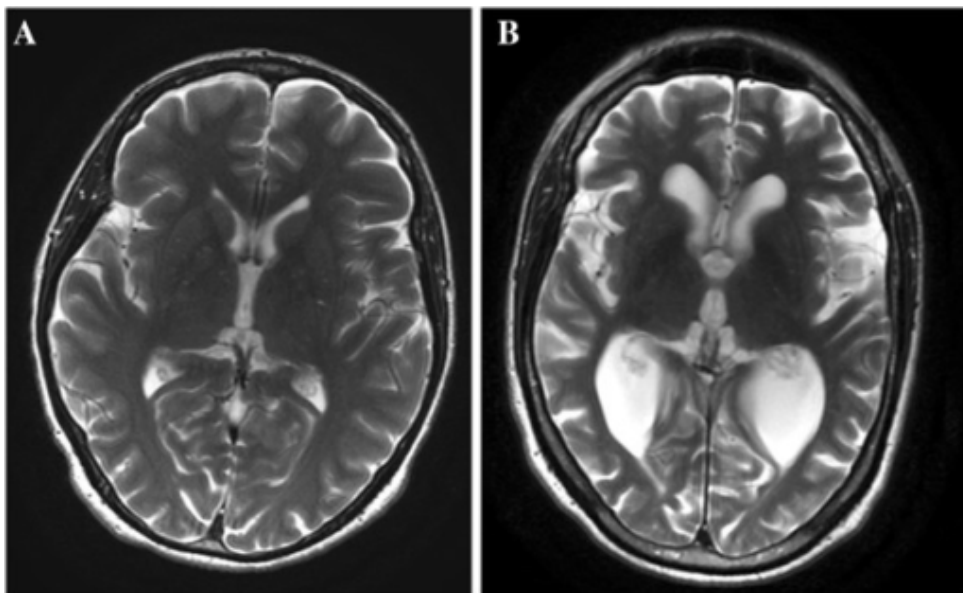


Figura 2.1: Risonanza Magnetica cerebrale. Le immagini assiali T2 mostrano atrofia cerebrale nel paziente con Alzheimer (A) e un cervello normale in un paziente di controllo della stessa età (B) (Cornelius et al., 2013)

### 2.1.3 MRI Funzionale e di Diffusione per l'analisi della connettività

Mentre la sMRI analizza la struttura anatomica, esistono altre modalità di MRI che permettono di indagare riguardo alla funzionalità e alla microstruttura cerebrale, rivelando spesso alterazioni persino precedenti all'atrofia visibile.

#### MRI di Diffusione

La MRI di Diffusione (DTI – *Diffusion Tensor Imaging*) sfrutta la sensibilità della Risonanza Magnetica al moto browniano: un movimento termico casuale e incessante delle molecole d'acqua; misurare statisticamente questo fenomeno microscopico dell'acqua permette di identificare la struttura dei tessuti biologici che ne ostacolano la diffusione. In un tessuto come la sostanza bianca, che è un tessuto fibroso composto da assoni mielinizzati, la diffusione dell'acqua è anisotropica, facilitata nella direzione delle fibre e ostacolata perpendicolarmente ad esse [14]. Nella patologia di Alzheimer, tuttavia, la mielina si degrada e gli assoni muoiono, riducendo questa direzionalità e permettendo al tensore di diffusione di rilevare parametri come la *Fractional Anisotropy* (FA) e la *Mean Diffusivity* (MD). FA e MD mostrano precocemente alterazioni in aree specifiche come il corpo calloso e il fascio del cingolo, a indicare una disconnessione tra le strutture cerebrali [16].

#### MRI Funzionale

La MRI Funzionale (fMRI) misura l'attività cerebrale rilevando variazioni nel flusso sanguigno e nell'ossigenazione dei tessuti, basandosi sul contrasto BOLD (*Blood Oxygenation Level Dependent*) e mostrando così quali aree comunicano fra loro [14]. Ciò è reso possibile dalle differenti proprietà magnetiche che hanno emoglobina ossigenata e deossigenata, perciò, quando un'area cerebrale è attiva, l'aumento del flusso sanguigno in quella zona modifica il segnale. Nello studio dell'AD è particolarmente utile e di interesse la *Resting State fMRI* (rs-fMRI), ovvero un'acquisizione fMRI eseguita con il paziente a riposo. Questo permette di analizzare le reti di connettività funzionale, evidenziando come la malattia comprometta precocemente il *Default Mode Network* (DMN), una rete neurale attiva durante il riposo e fondamentale per la memoria episodica [16].

## 2.2 La Tomografia a Emissione di Positroni

Se da un lato la MRI mostra la morfologia dell'encefalo con dettagli ad alta risoluzione, dall'altro la Tomografia a Emissione di Positroni (PET – *Positron Emission Tomography*) permette di visualizzare i processi metabolici e le funzioni chimiche *in vivo* [14]. Questo è molto importante nel contesto della diagnosi precoce di Alzheimer, in quanto permette di misurare alterazioni fisiopatologiche come l'ipometabolismo o l'accumulo di proteine tossiche, spesso presenti già prima che si verifichi un'atrofia strutturale misurabile tramite sMRI.

### 2.2.1 Principi fisici e traccianti

L'acquisizione delle immagini PET si basa su radiofarmaci marcati con isotopi che emettono positroni e che quindi hanno un decadimento di tipo  $\beta^+$ , come per esempio il Fluoro-18 ( $^{18}\text{F}$ ) o il Carbonio-11 ( $^{11}\text{C}$ ). Quando l'isotopo contenuto nel farmaco decade nel corpo, emette un positrone che viaggia per una breve distanza, solitamente pochi millimetri, prima di incontrare un elettrone nel tessuto circostante; l'interazione tra queste due particelle causa un processo di annichilazione, che, convertendo le masse in energia, genera due fotoni gamma ( $\gamma$ ) da 511 keV emessi a circa  $180^\circ$  l'uno dall'altro [14]. Il rilevamento di questo processo non è affidato a collimatori meccanici, ma si basa sulla collimazione elettronica, una tecnica che sfrutta il rilevamento in coincidenza: l'evento viene registrato soltanto se due rilevatori posti sui lati opposti dell'anello rilevano un fotone in una finestra temporale molto ristretta, solitamente dell'ordine di pochi nanosecondi. La linea che congiunge questi rilevatori, chiamata *Line of Response* o LOR, identifica in quale direzione è avvenuta l'annichilazione; in seguito, algoritmi iterativi ricostruiscono l'insieme delle LOR per generare l'immagine della distribuzione del tracciante [14].

### 2.2.2 FDG-PET per la valutazione dell'ipometabolismo

Il tracciante più utilizzato in ambito neurologico è il fluorodesossiglucosio marcato con fluoro-18 ( $^{18}\text{F}$ -FDG), in quanto chimicamente simile al glucosio, che rappresenta la fonte principale di energia utilizzata dal cervello; il tracciante, per questo motivo, si accumula nelle aree a maggiore attività sinaptica e neuronale, che richiedono più energia. Nella malattia di Alzheimer, l'FDG-PET mostra pattern di ipometabolismo, quindi di ridotto consumo di glucosio, indicando sinapsi disfunzionali oppure morte neuronale [8]. L'ipometabolismo colpisce aree specifiche, come le regioni temporo-parietali, il cingolo posteriore e il precuneo, spesso mostrando simmetria come visibile in Figura 2.2. Questo cambiamento metabolico è molto utile a prevedere il passaggio da MCI a demenza conclamata, anticipando i cambiamenti nella morfologia [18].

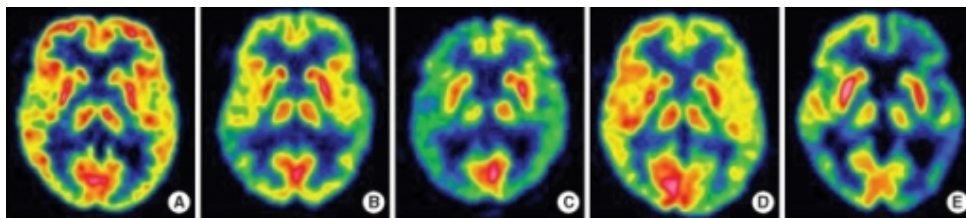


Figura 2.2: Confronto di immagini FDG-PET tra un controllo CN (A) e quattro pazienti affetti da Alzheimer ad esordio precoce (B-E). Le immagini FDG-PET mostrano diversi gradi di ipometabolismo nella corteccia frontale e temporo-parietale (Park et al., 2008)

### 2.2.3 Amyloid-PET e Tau-PET per i biomarker specifici

Negli ultimi anni sono stati sviluppati nuovi traccianti specifici, che permettono di mostrare direttamente i due segni patologici distintivi dell'AD: le placche di beta-Amiloide e i grovigli neurofibrillari di proteina Tau. La Amyloid-PET utilizza traccianti come il  $^{11}\text{C}$ -PiB (*Pittsburgh Compound B*) o i composti fluorurati  $^{18}\text{F}$ -Florbetapir e  $^{18}\text{F}$ -Flutemetamol che si legano alle placche di Amiloide. Come descritto nel primo capitolo, la positività all'Amiloide è il primo evento nella catena della patologia di Alzheimer e può essere rilevata anche 15-20 anni prima che si presentino i primi sintomi clinici [7]. La Tau-PET, invece, utilizza leganti specifici, come il  $^{18}\text{F}$ -Flortaucipir, che si legano ai grovigli di proteina Tau. A differenza dell'Amiloide, che si deposita in maniera diffusa, la distribuzione dei grovigli di Tau è strettamente legata alla neuro-degenerazione e ai danni cognitivi specifici, rappresentando un biomarcatore molto preciso [7].

Queste specifiche modalità rappresentano quindi le "etichette biologiche" necessarie ad addestrare i modelli di AI a distinguere l'Alzheimer da altre forme di demenza.

## 2.3 La Tomografia Computerizzata e il suo ruolo

Sebbene sia la Risonanza Magnetica il gold standard per il neuroimaging nell'AD, la Tomografia Computerizzata (CT – *Computed Tomography*) svolge comunque un ruolo ben definito. Infatti, viene utilizzata soprattutto nelle fasi di screening o in casi in cui non si può ricorrere alla MRI [20]. A differenza di quest'ultima, la CT utilizza radiazioni ionizzanti, sfruttando le proprietà di attenuazione di un fascio di raggi X che passa attraverso il corpo. Attraverso ciò, la CT permette di generare immagini tomografiche basate sulla diversa densità dei tessuti. Le componenti principali di uno scanner CT sono due: un tubo radiogeno e un arco di rivelatori posizionati su un *gantry* che ruotano attorno al paziente; il fascio di raggi X attraversa quindi il paziente passando per la parte del corpo interessata e subisce un'attenuazione descrivibile dalla legge di Lambert-Beer [21]:

$$I = I_0 e^{-\int \mu(x,y) dl} \quad (2.2)$$

dove  $I_0$  è l'intensità del fascio incidente,  $I$  è l'intensità rilevata e  $\mu(x, y)$  è il coefficiente di attenuazione lineare locale del tessuto in un punto individuato dalle coordinate  $(x, y)$ .

I dati ottenuti e acquisiti dai rivelatori riguardo l'attenuazione del fascio vengono poi elaborati tramite algoritmi, come la *Filtered Back Projection* (FBP), per ottenere una matrice di pixel. Ognuno di questi pixel è espresso in Unità Hounsfield (HU), una scala normalizzata in cui 0 HU rappresenta il valore dell'acqua e -1000 HU invece rappresenta il valore associato all'aria; utilizzando questa scala è quindi possibile caratterizzare la densità dei diversi tessuti [14].

### 2.3.1 Applicazione clinica nell'Alzheimer

Storicamente, il ruolo dell'imaging strutturale come la CT nella diagnosi di demenze e AD era limitato all'esclusione di complicanze neurochirurgiche come ematomi subdurali, neoplasie cerebrali, idrocefalo ecc. L'approccio che si segue attualmente, invece, non si basa solamente sull'esclusione, ma anche sulla ricerca di caratteristiche che possano supportare la diagnosi [20]. Specificamente nell'ambito dell'Alzheimer, la CT è molto efficace e utile per valutare patologie concomitanti, grazie alla sua capacità di vedere alterazioni nella sostanza bianca, come può capitare per l'ipodensità; questo permette di discriminare la diagnosi tra Alzheimer puro, demenza vascolare o altre forme miste [20].

## 2.4 Il pre-processing delle immagini: dalla strumentazione al dato analizzabile

Una volta acquisite le neuroimmagini, sono disponibili dati grezzi che, però, raramente sono adatti ad essere analizzati direttamente; infatti, queste possono presentare grandi differenze e variabilità tra un soggetto e un altro, dovute a differenze nell'hardware, nei protocolli d'acquisizione e nel posizionamento del paziente. Per trasformare questi dati eterogenei in dati normalizzati e standardizzati è necessario il pre-processing: un insieme di correzione, rimozione dei difetti e normalizzazione che ha l'obiettivo di massimizzare il rapporto segnale-rumore (SNR) e garantire che le differenze che potrebbe rilevare un modello di classificazione siano causate da variazioni biologiche reali, piuttosto che da artefatti tecnici. Nel contesto della ricerca sull'AD, sono state definite pipeline di pre-processing rigorose e standardizzate da applicare prima di qualsiasi analisi morfometrica, grazie a iniziative come l'*Alzheimer's Disease Neuroimaging Initiative* (ADNI) [22]. Senza questa prima fase, non si potrebbe ritenere affidabile nessun classificatore basato su Machine Learning o Deep Learning, a causa del fenomeno "garbage in, garbage out": la qualità delle immagini di input determina la validità dell'output [23].

### 2.4.1 Rimozione degli artefatti e Skull Stripping

Per analizzare correttamente il cervello, il primo passo nella catena di pre-processing è la rimozione di tutti i tessuti non cerebrali; questo processo è noto come *skull stripping*. Le immagini non ancora elaborate, infatti, includono ossa del cranio, muscoli, occhi ecc. Queste strutture non sono utili per la diagnosi di Alzheimer e possono persino introdurre rumore durante l'addestramento di una rete neurale. Uno degli algoritmi più utilizzati per lo *skull stripping* è il *Brain Extraction Tool* (BET), che sfrutta un modello deformabile utile a definire la superficie del cervello. Il BET è un algoritmo veloce, che non richiede nemmeno pre-processing ulteriore prima di essere utilizzato [24]. In questa fase, bisogna essere precisi: rimuovere eccessivamente potrebbe

eliminare alcune porzioni della corteccia cerebrale, al contrario, rimuovere insufficientemente potrebbe lasciare residui di tessuti non cerebrali, alterando i passaggi successivi.

## 2.4.2 Correzione delle disomogeneità

Esiste inoltre un artefatto che si può incontrare comunemente nelle immagini MRI, ovvero la disomogeneità del campo di intensità, chiamata anche *bias field*. Questo è un fenomeno rappresentato da una variazione a bassa frequenza dell'intensità del segnale e che può essere causata da imperfezioni nel campo magnetico o dalla posizione del paziente nello scanner [25]. A causa di questo artefatto, un tessuto può apparire con livelli di grigio significativamente differenti in diverse regioni del cervello, come viene evidenziato in Figura 2.3. Questo è un problema che rende difficile la distinzione automatica dei tessuti e che quindi deve essere corretto; per ridurlo, si è cercato di sviluppare algoritmi di correzione specifici. Attualmente, lo standard di riferimento è l'algoritmo N4ITK, che punta a migliorare il precedente N3 (*Non-parametric Non-uniform intensity Normalization*). È stato dimostrato che applicare l'algoritmo N4ITK riduce notevolmente la variazione dell'intensità tra i tessuti uguali, permettendo di avere dati più coerenti e rendendo più semplice estrarre feature affidabili per analisi statistiche e per il Deep Learning [26].

## 2.4.3 Registrazione e Normalizzazione

Le differenze fisiologiche di forma e dimensione tra cervelli di diversi soggetti rende difficile confrontarli voxel per voxel. Per questo è prima necessario allinearli in un sistema di coordinate comune; questo processo è composto da due fasi: registrazione e normalizzazione. La registrazione allinea le immagini tramite sole trasformazioni lineari, che possono essere traslazione, rotazione e scaling; in seguito, la normalizzazione spaziale deforma l'immagine per adattarla a un modello di riferimento [28]. Lo standard utilizzato come modello è solitamente il template MNI (*Montreal Neurological Institute*). Per eseguire la normalizzazione spaziale si utilizzano trasformazioni non lineari complesse, a differenza della registrazione, per correggere le differenze locali e poter quindi identificare l'atrofia specifica tipica dell'Alzheimer; questo assicura che una coordinata specifica  $(x, y, z)$  corrisponda alla stessa struttura anatomica per tutte le immagini nel dataset. Oltre alla normalizzazione spaziale, è importante eseguire anche una normalizzazione di intensità dei voxel [28]; infatti, a differenza della CT, in cui i valori dell'intensità hanno un significato fisico assoluto grazie alle Unità Hounsfield, nella MRI questi dipendono dal guadagno dello scanner e dal protocollo utilizzato; di conseguenza, uniformare questi valori è importante per garantire che le variazioni rilevate dagli algoritmi AI siano dovute a reali differenze tra i tessuti e non causate da errori strumentali.

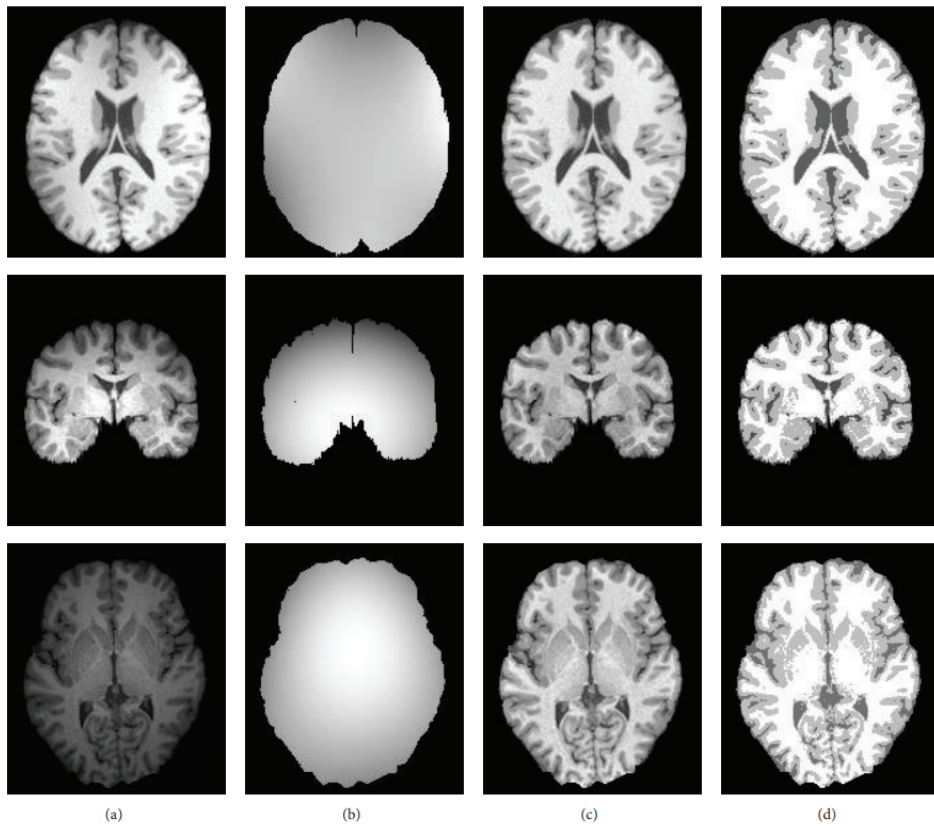


Figura 2.3: *Esempio di stima del bias field e risultati delle correzioni in immagini MRI. (a) Immagine originale; (b) Stima del bias field; (c) Immagini con bias field rimosso; (d) Immagini segmentate* (Cong et al., 2016)

#### 2.4.4 Software di riferimento nella ricerca

Esistono diversi pacchetti software accademici open-source, che rappresentano il gold standard nella ricerca neuroscientifica e che rendono possibile implementare le pipeline di pre-processing descritte in precedenza. Alcuni di questi sono:

- **SPM** (Statistical Parametric Mapping): è un pacchetto software sviluppato presso l'UCL di Londra. Comprende una pipeline per la *Voxel-Based Morphometry*, completa di pre-processing, smoothing e analisi statistica [29].
- **FSL** (FMRIB Software Library): sviluppato dall'Università di Oxford, è un pacchetto gratuito per uso non commerciale. Include strumenti come BET per lo *skull stripping* [30].
- **FreeSurfer**: è un software specializzato nella ricostruzione della superficie corticale e nella segmentazione automatica delle strutture sottocorticali. È particolarmente rinomato per la sua capacità di misurare lo spessore corticale e il volume di alcune strutture specifiche con elevata precisione [31].

Questi strumenti sono spesso integrati nelle pipeline automatizzate per gestire i grandi volumi di dati richiesti per addestrare i modelli di Deep Learning.

## Capitolo 3

# Metodologie di Intelligenza Artificiale per la Classificazione

L'analisi dei dati biomedici è in continua evoluzione; tradizionalmente, la ricerca si è sempre affidata alla statistica per testare ipotesi e per individuare differenze tra gruppi di popolazioni. Questo approccio è utile a comprendere i meccanismi generali delle patologie, tuttavia è anche limitato quando l'obiettivo passa dalla descrizione di campioni alla predizione sul singolo individuo, che è una necessità per la medicina di precisione. Il Machine Learning (ML) è nato proprio per risolvere questo problema.

### 3.1 Approcci di Machine Learning tradizionale

Per definizione, il Machine Learning è un algoritmo di apprendimento automatico – o un sistema – capace di migliorare le sue prestazioni nello svolgere un determinato compito tramite l'esperienza, rendendo essenzialmente un computer in grado di comprendere il mondo in termini di "gerarchie di concetti" [32]. Il processo classico degli approcci tradizionali di ML (o *shallow learning*) segue la metodologia dell'apprendimento supervisionato ed è strutturato in due fasi:

1. **Feature Engineering:** l'estrazione manuale (o semi-automatica) di caratteristiche rilevanti delle immagini;
2. **Classificazione:** l'addestramento di un algoritmo su queste caratteristiche (feature) per prevedere la classe di appartenenza del soggetto [33].

L'addestramento prevede che le informazioni di interesse siano codificate e che vengano fornite poi al modello sotto forma di:

- **Features:** variabili di input che rappresentano le caratteristiche del dato;
- **Labels:** variabili di output che indicano la classe di appartenenza (per esempio "Soggetto Sano" o "AD") [34].

Attraverso algoritmi di apprendimento, il sistema identifica la funzione che mappa le feature sulle rispettive label. L'obiettivo finale è ottenere un modello capace di generalizzare, ovvero mantenere un'elevata accuratezza predittiva anche quando viene valutato su nuovi set di dati mai analizzati; inoltre, a differenza degli operatori umani, il sistema è in grado di processare enormi quantità di dati senza cali di prestazioni da affaticamento cognitivo o variabilità imputabile a interpretazioni soggettive [34].

### 3.1.1 Estrazione manuale delle feature

Applicare gli algoritmi di Machine Learning direttamente sui dati di neuroimaging grezzi è reso difficile a causa di uno specifico problema, noto come *curse of dimensionality* (letteralmente "maledizione della dimensionalità"), dovuto alle caratteristiche intrinseche dei dataset di bioimmagini. Questi dataset, infatti, sono per loro natura ad alta dimensionalità: un singolo volume cerebrale può contenere infatti milioni di voxel, mentre il numero dei soggetti disponibili per l'addestramento ( $n$ ) è inferiore di diversi ordini di grandezza. Questa condizione in cui il numero delle variabili ( $p$ ) supera di gran lunga il numero delle osservazioni ( $p \gg n$ ) rende lo spazio delle feature estremamente sparso. Utilizzare tali dati senza un pre-processing adeguato compromette la capacità del modello di generalizzare sui nuovi pazienti [35]. Per ridurre questo problema ci si affida al processo di Feature Engineering: una fase di preparazione in cui i dati grezzi vengono trasformati e rappresentati in maniera più compatta e informativa, sfruttando le conoscenze specifiche degli operatori (in questo caso anatomiche, neurologiche e cliniche) e applicandole ai dati per facilitare l'algoritmo di apprendimento [33]. L'obiettivo del Feature Engineering è quindi massimizzare l'accuratezza predittiva del dataset riducendo la dimensionalità; per farlo è necessario lavorare su due tipologie di feature:

- **Feature Irrilevanti:** variabili che non possiedono alcuna correlazione con la diagnosi; vanno tassativamente escluse poiché introducono rumore nel modello [35];
- **Feature Ridondanti:** variabili che portano informazioni simili tra loro. Nonostante la statistica tradizionale di norma suggerisca di rimuoverle per alleggerire il carico computazionale, è altrettanto vero che, nell'addestramento di modelli complessi, risulta utile mantenere una certa ridondanza per aumentarne la robustezza [33]; in presenza di rumore, infatti, avere variabili ridondanti che contengono la stessa informazione evita che la corruzione di una singola feature comprometta drasticamente l'accuratezza predittiva.

Una delle sfide principali in questa fase di Feature Engineering è il bilanciamento tra quello che viene chiamato *overfitting* e *underfitting*. Solitamente si punta a ottenere una condizione intermedia definita *optimum fitting* [35]:

- **Overfitting (Alta Varianza):** è una condizione che si verifica quando il modello è troppo complesso. In questo caso, il rumore statistico viene memorizzato nel dataset e il modello fallisce nel diagnosticare nuovi pazienti;

- **Underfitting** (Alto Bias): si verifica quando il modello è, al contrario, troppo semplice o dispone di un numero insufficiente di feature. Questo comporta una riduzione della capacità del modello di comprendere la complessità e le sfumature della patologia.

### 3.1.2 Support Vector Machines

Tra gli algoritmi di Machine Learning più comuni nella diagnosi di Alzheimer, le Support Vector Machines (SVM) hanno rappresentato a lungo il gold standard per le analisi morfometriche e tuttora sono una solida *baseline* di confronto per validare i modelli di Deep Learning più moderni. Il fondamento teorico delle SVM, introdotto da Cortes e Vapnik nel 1995, si basa sul minimizzare il rischio strutturale, ovvero sul ricercare il miglior compromesso matematico tra l'accuratezza sui dati e la prevenzione dell'*overfitting* [36]. Geometricamente, questo si traduce nell'obiettivo di individuare un iperpiano di separazione che massimizzi le differenze tra due classi (nel caso specifico, tra soggetti sani e pazienti AD). I punti che giacciono esattamente sul margine di maggior separazione vengono chiamati "Vettori di Supporto" (da cui il nome dell'algoritmo) e sono gli unici elementi del dataset che determinano la definizione del modello [36]. Matematicamente, dato un set di dati  $\{(x_i, y_i)\}$ , l'algoritmo risolve un problema di ottimizzazione quadratica per trovare il vettore dei pesi  $w$  e il *bias*  $b$  che definiscono l'iperpiano. La funzione di decisione [32, 36] per classificare un nuovo campione è rappresentata dall'equazione:

$$f(x) = \text{sign}(w \cdot x + b) \quad (3.1)$$

Tuttavia, i dati biomedici raramente sono separabili in modo lineare nel loro spazio originale, per cui, per aggirare questa limitazione, bisogna definire una *kernel function*: una funzione matematica del tipo  $K(x_i, x_j)$  che proietta i dati in ingresso in uno spazio di dimensione superiore, in cui è poi possibile trovare un iperpiano di separazione lineare senza doverne calcolare le coordinate in modo esplicito.

L'efficacia clinica di questo approccio è stata dimostrata dallo studio pionieristico di Magnin et al., in cui gli autori hanno applicato una SVM lineare all'analisi di immagini MRI *whole-brain* [37]. Questo approccio ha permesso di identificare pattern di atrofia neurodegenerativa, con un'accuratezza diagnostica del 94.3% nella discriminazione tra pazienti AD e controlli sani, a dimostrazione di come questo algoritmo sia capace di distinguere differenze abbastanza sottili da sfuggire alla semplice ispezione visiva [37]. Nonostante la grande importanza delle SVM, la letteratura recente mostra come stia avvenendo una rapida transizione alle Reti Neurali Convoluzionali (CNN), grazie alle loro superiori performance diagnostiche; tuttavia, questo vantaggio comporta un carico computazionale elevato e una minore interpretabilità rispetto ai metodi tradizionali [38]. Queste sfide, abbinate a queste grandi potenzialità, definiscono il campo del Deep Learning.

## 3.2 Approcci di Deep Learning

Il Deep Learning (DL) è un sottoinsieme del Machine Learning in grande crescita che ha rivoluzionato l'analisi delle immagini biomediche. A differenza degli algoritmi tradizionali che richiedono un'estrazione manuale delle feature e Feature Engineering, il DL automatizza questo processo utilizzando le Reti Neurali Artificiali (ANN) profonde; queste reti sono strutturate come una composizione di funzioni differenti, il cui modello complessivo è rappresentato da una catena di funzioni associate ai diversi strati (*layers*) [32]:

$$f(\mathbf{x}) = f^{(n)}(\dots f^{(2)}(f^{(1)}(\mathbf{x}))\dots) \quad (3.2)$$

dove  $f^{(1)}$  è il primo livello della rete,  $f^{(2)}$  il secondo e così via. L'unità fondamentale di queste reti è il neurone artificiale, o perceptrone. Il suo funzionamento è analogo a quello di un neurone biologico, che riceve segnali elettrici dai dendriti e, se lo stimolo è sufficiente, trasmette un impulso lungo l'assone. Allo stesso modo, un neurone artificiale svolge tre operazioni matematiche [32]:

1. Riceve dei valori in input ( $\mathbf{x}$ );
2. Li moltiplica per dei pesi ( $\mathbf{W}^\top$ ) che rappresentano l'importanza di ogni collegamento;
3. Somma i risultati e aggiunge un valore di soglia chiamato *bias* ( $\mathbf{c}$ )

La formula che descrive questo processo è quindi [32]:

$$\mathbf{h} = g(\mathbf{W}^\top \mathbf{x} + \mathbf{c}) \quad (3.3)$$

Dove  $g$  è la funzione di attivazione (non lineare). Questa funzione è importante perché decide se il neurone deve attivarsi o meno, permettendo alla rete di apprendere relazioni complesse e non lineari, molto comuni nei dati biomedici.

Un singolo neurone non è molto potente; la forza del Deep Learning è data dal collegamento di migliaia di questi neuroni in una struttura a strati:

- **Input Layer:** riceve i dati grezzi;
- **Hidden Layers:** sono gli strati nascosti intermedi. Più ce ne sono, più la rete è "profonda" (*Deep*); ognuno di questi strati elabora le informazioni ricevute da quello precedente, estraendo caratteristiche sempre più astratte;
- **Output Layer:** fornisce il risultato finale.

L'addestramento della rete avviene tramite un algoritmo di *back-propagation* ("retropropagazione"): la rete confronta la sua previsione con l'output reale; se sbaglia, l'errore viene calcolato e

propagato nella direzione opposta per aggiustare poco a poco tutti i pesi  $W^T$  della rete, finché non diventa minimo [32].

### 3.2.1 Le Reti Neurali Convoluzionali per l'analisi di immagini

Nell'analisi delle bioimmagini, le Reti Neurali Convoluzionali (CNN) rappresentano oggi lo standard di riferimento tra le diverse architetture di Deep Learning. Queste reti risolvono il problema che deriva dal gestire dati ad alta dimensionalità, come possono essere le immagini MRI, sfruttando la topologia dell'immagine. Per farlo, le CNN si basano su tre concetti chiave: connettività locale, condivisione dei pesi e invarianza alle traslazioni [39]. L'architettura tipica alterna tre tipi di livelli:

1. **Livelli Convoluzionali** (*Convolutional Layers*): sono il cuore della rete; invece di una calcolare una matrice completa di pesi, la rete applica un filtro (o *kernel*) che fa scorrere sull'immagine. Matematicamente, l'operazione di convoluzione discreta tra un input bidimensionale  $I$  e un kernel  $K$  è definita come:

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n) \quad (3.4)$$

Concretamente, però, nelle librerie di Deep Learning l'operazione implementata è solitamente la cross-correlazione discreta, che per un input bidimensionale  $I$  e un *kernel*  $K$ , sfruttando la commutatività della convoluzione, è definita come:

$$S(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i + m, j + n)K(m, n) \quad (3.5)$$

Questo processo genera delle *feature maps*, mappe che mostrano la presenza di pattern specifici in ogni punto dell'immagine;

2. **Livelli di Pooling**: eseguono un *down-sampling* delle *feature maps*, ovvero le sottocampionano. Così facendo, si riduce la dimensionalità spaziale e si rende la rappresentazione che ne deriva robusta a piccole traslazioni o a naturali distorsioni anatomiche del cervello del paziente;
3. **Livelli Fully Connected** (FC): posti alla fine della rete. I livelli FC "appiattiscono" le feature estratte in un singolo vettore monodimensionale, sul quale viene poi eseguita la classificazione finale [32].

La maggior parte delle architetture classiche opera di base su immagini in 2D, considerando il volume delle immagini MRI come una sequenza di "fette". Questo comporta però una perdita di informazioni lungo l'asse della profondità  $z$ , che limita la capacità del modello di ricostruire la morfologia reale di strutture anatomiche complesse [38]. Per superare questo limite dato dalla

perdita di continuità spaziale, la ricerca si sta orientando verso l'uso di 3D-CNN strutturate come nell'esempio in Figura 3.1, che utilizzano *kernel* volumetrici ( $x, y, z$ ) per elaborare l'intera scansione in una volta. Questo approccio è superiore nel riscontrare pattern di atrofia, ma comporta compromessi importanti: infatti, l'aumento esponenziale dei parametri da addestrare richiede risorse di calcolo (come GPU) molto più elevate e dataset più grandi per evitare l'*overfitting*, motivo per cui le architetture 2D sono ancora una scelta molto valida e diffusa in contesti con risorse limitate [38].

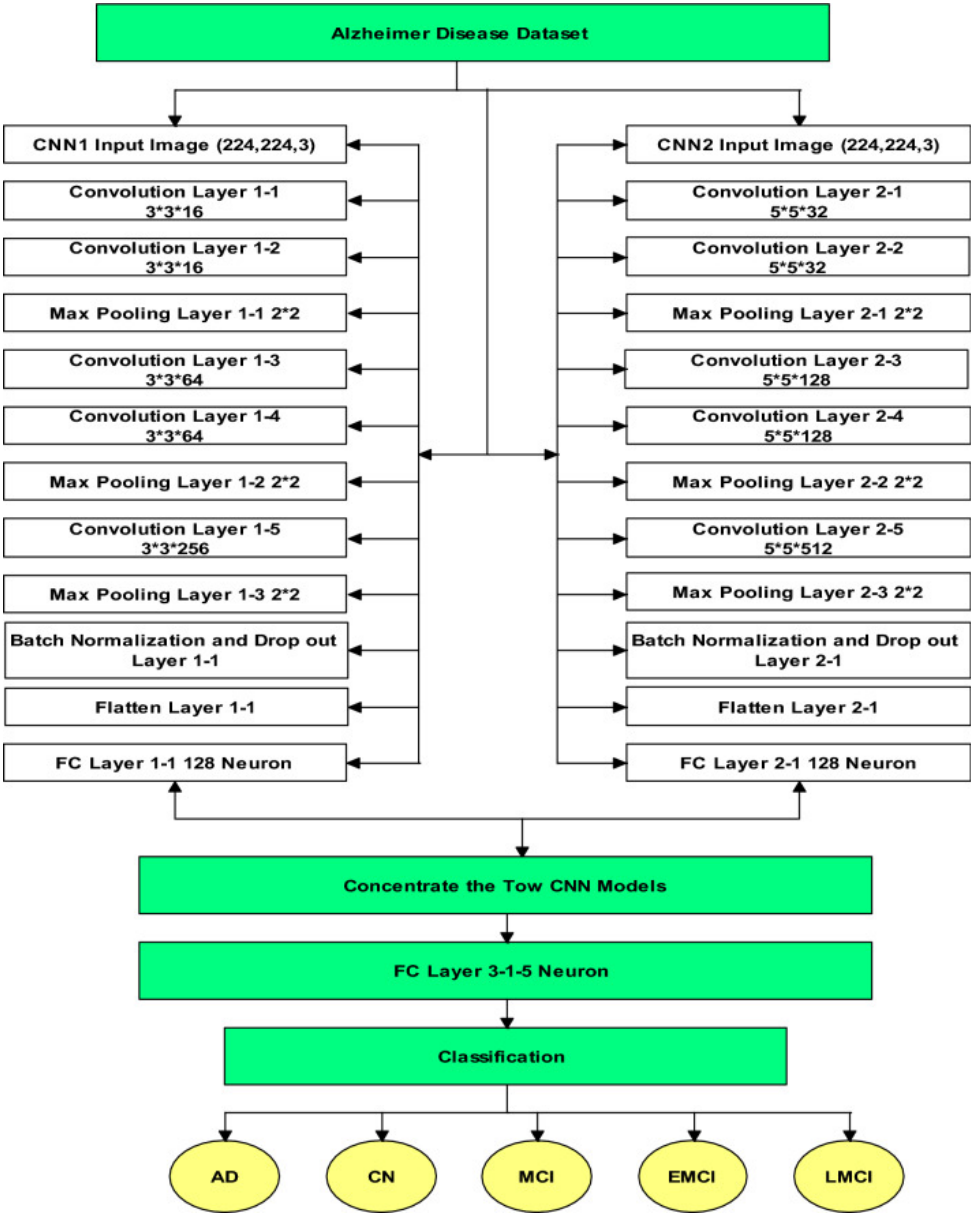


Figura 3.1: Un esempio di architettura CNN per la rilevazione e la classificazione dell'Alzheimer (El-Assy et al., 2024)

### 3.2.2 Architetture comuni e Transfer Learning

Nello sviluppo di sistemi per la diagnosi di Alzheimer, le reti neurali raramente vengono sviluppate a partire da zero. Al contrario, di prassi si prevede l'utilizzo di architetture CNN standard che hanno dimostrato prestazioni eccellenti in compiti di *computer vision* generale. Tra le architetture più citate figurano [41]:

- **VGGNet** (VGG-16, VGG-19): introdotta da Simonyan e Zisserman, è un'architettura caratterizzata da una grande semplicità strutturale e profondità. Utilizzando solamente filtri convoluzionali  $3 \times 3$ , può apprendere feature con pattern gerarchici molto complessi [42].
- **ResNet** (*Residual Network*): sviluppata da He et al. per risolvere il problema della degradazione del gradiente (*vanishing gradient*) nelle reti profonde. L'innovazione più importante è l'introduzione delle connessioni residue (o *skip connections*), che permettono al segnale di bypassare uno o più strati. Questo meccanismo di apprendimento consente di addestrare reti con centinaia o migliaia di *layer* senza perdita di performance [43].
- **Inception**: proposta da Szegedy et al., team di ricerca di Google, inizialmente con il nome di GoogLeNet, poi diventato Inception v1. Si basa sul modulo "Inception", un blocco che applica diverse operazioni di convoluzione, utilizzando *kernel* di dimensioni  $1 \times 1$ ,  $3 \times 3$  e  $5 \times 5$  sullo stesso input. Questo approccio permette alla rete di catturare diversi dettagli a diverse risoluzioni spaziali contemporaneamente, riducendo allo stesso tempo il numero di parametri computazionali [44].

Tuttavia, addestrare da zero queste architetture presenta un ostacolo: la scarsità di dati biomedici etichettati. I dataset di neuroimaging come ADNI, infatti, contano solo poche migliaia di scansioni; addestrare una rete con milioni di parametri su un dataset così ridotto porterebbe inevitabilmente all'*overfitting* e porterebbe la rete a memorizzare i pochi esempi di training senza però riuscire a generalizzare sui nuovi pazienti [45]. La tecnica principale a cui si ricorre per superare questo limite è il *Transfer Learning*, che consiste nell'utilizzare una rete pre-addestrata su un dominio di origine molto vasto – come per esempio il dataset ImageNet, che contiene milioni di immagini naturali [46] – e riutilizzare i pesi appresi per il compito target. I primi livelli di una CNN apprendono infatti feature "di basso livello" (contorni, curve, texture ecc.), che sono universali e quindi trasferibili tra domini diversi [45]. Secondo lo studio di Tajbakhsh et al., esistono due strategie principali di *Transfer Learning* applicabili all'imaging biomedico [45]:

1. **Feature Extraction**: è una strategia in cui la rete pre-addestrata viene utilizzata come un estrattore di feature; quelle estratte dall'ultimo livello vengono poi usate per addestrare un classificatore tradizionale.

2. **Fine-Tuning**: in questa, i pesi della rete pre-addestrata vengono utilizzati come inizializzazione, ma vengono in seguito aggiornati tramite *back-propagation* sul dataset specifico che si vuole classificare.

### 3.3 Metriche di valutazione dei modelli

Per quantificare l'efficacia delle diagnosi dei modelli di Machine Learning e Deep Learning descritti precedentemente, la semplice accuratezza può essere fuorviante, soprattutto in ambito medico dove i dataset sono spesso sbilanciati, per esempio includendo molti più soggetti sani che malati. Servono quindi metriche di valutazione definite in modo rigoroso per valutare i modelli ML e DL.

#### 3.3.1 Metriche standard: Accuratezza, Sensitività, Specificità, Precisione, Curva ROC/AUC

La base per il calcolo di tutte le metriche prestazionali è la Matrice di Confusione. Questa tabella  $2 \times 2$  confronta le previsioni del modello con le etichette reali, classificando ogni soggetto in quattro categorie [47]:

- **Veri Positivi (TP)**: pazienti malati correttamente identificati come malati;
- **Veri Negativi (TN)**: soggetti sani correttamente identificati come sani;
- **Falsi Positivi (FP)**: soggetti sani erroneamente classificati come malati;
- **Falsi Negativi (FN)**: pazienti malati erroneamente classificati come sani.

A partire da questi valori, vengono definite le seguenti metriche scalari:

1. **Accuratezza (Accuracy)**: rappresenta la frazione di previsioni corrette sul totale dei casi esaminati.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3.6)$$

2. **Sensibilità (Recall o Sensitivity)**: misura la capacità del modello di rilevare la patologia quando questa è presente; in ambito clinico, è importante per minimizzare i falsi negativi.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3.7)$$

3. **Specificità (Specificity)**: misura la capacità del modello di identificare i soggetti sani, evitando i falsi allarmi.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (3.8)$$

4. **Precisione** (*Precision*): misura l'affidabilità del modello quando predice la classe positiva.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3.9)$$

5.  $F_\beta$ -**score**: utilizzato per sintetizzare Precisione e Sensibilità in un unico indice.

$$F_\beta = \frac{(\beta^2 + 1)\text{TP}}{(\beta^2 + 1)\text{TP} + \beta^2\text{FN} + \text{FP}} \quad (3.10)$$

Dove  $\beta$  è un parametro che determina il peso relativo della Sensibilità rispetto alla Precisione. Per avere uguale importanza tra le due metriche, si pone  $\beta = 1$ , ottenendo l' $F_1$ -**score**, che rappresenta la media armonica tra Precisione e Sensibilità:

$$F_1 = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FN} + \text{FP}} \quad (3.11)$$

L' $F_1$ -score è molto importante nei dataset biomedici sbilanciati, poiché penalizza i modelli che ottengono un'alta accuratezza semplicemente ignorando la classe minoritaria (ad esempio, su un dataset di 90 soggetti sani e 10 pazienti malati, un modello che classifica tutti i soggetti come sani otterrebbe un'accuratezza del 90%, anche se di fatto risulterebbe un classificatore clinicamente inefficace).

Oltre a queste metriche, lo standard per valutare un classificatore binario è la curva ROC (*Receiver Operating Characteristic*), un grafico che illustra la capacità diagnostica di un sistema al variare della soglia di discriminazione, evidenziando la relazione tra il Tasso di Veri Positivi (TPR – *True Positive Rate*, equivalente alla Sensibilità) sull'asse  $y$  e il Tasso di Falsi Positivi (FPR – *False Positive Rate*, calcolato come  $1 - \text{Specificity}$ ) sull'asse  $x$ . Per confrontare diversi modelli con un singolo valore numerico, si utilizza la *Area Under the Curve* (AUC); essa rappresenta la probabilità che il classificatore assegni un punteggio più alto a un soggetto positivo scelto casualmente rispetto a un soggetto negativo scelto casualmente [48]. Un valore di AUC pari a 0.5 indica un classificatore casuale, con nessuna capacità discriminante, mentre un valore pari a 1.0 indica un classificatore perfetto.

### 3.3.2 Strategie di validazione robusta: Cross-Validation e prevenzione del Data Leakage

Poiché i campioni disponibili nei dataset biomedici (come ADNI) sono spesso limitati in numero, dividere i dati in *training set* e *test set* può portare a performance inaffidabili e fortemente dipendenti dalla casualità della partizione. Per garantire che i risultati siano statisticamente robusti, sono necessarie strategie di validazione rigorose, prevenendo allo stesso tempo fenomeni di contaminazione dei dati.

## **k-Fold Cross-Validation**

La tecnica standard per avere una stima imparziale delle prestazioni è la *k-Fold Cross Validation*, analizzata formalmente da Kohavi nel 1995. In questa metodologia, il dataset viene suddiviso in  $k$  parti uguali chiamate *folds* e si ripete il processo di addestramento per  $k$  volte: in ogni iterazione, un *fold* differente viene utilizzato come set di test, mentre i restanti  $k - 1$  formano il set di training. La performance finale del modello viene poi calcolata come la media delle metriche ottenute nelle  $k$  iterazioni. Valori moderati di  $k$  (10-20) rappresentano il compromesso ottimale tra bias e varianza della stima [49].

## **Prevenzione del Data Leakage**

Nello sviluppo di algoritmi per l'analisi di immagini biomediche, uno dei problemi più critici è il *Data Leakage* (letteralmente "fuga di informazioni"). Un errore frequente negli studi sull'Alzheimer che usano 2D-CNN, come evidenziato da Wen et al. nel 2020, consiste nell'effettuare la divisione tra training set e test set a livello di singola fetta piuttosto che a livello del singolo paziente. Se le sezioni dello stesso paziente finiscono sia nel set di addestramento che nel set di test, la rete neurale rischia di apprendere l'anatomia specifica del soggetto piuttosto che le caratteristiche fisiopatologiche della malattia, portando a una sovrastima delle prestazioni. Gli autori stessi hanno sottolineato la gravità di questo fenomeno, rilevando che più della metà degli studi da loro esaminati risultava affetto da Data Leakage [50]. Per evitare ciò, è fondamentale avere immagini indipendenti a livello di soggetto: tutte le immagini appartenenti allo stesso paziente devono essere usate esclusivamente in uno dei set, così da essere sicuri che il modello venga valutato su anatomie mai viste in fase di addestramento [50].

# Capitolo 4

## Analisi dello Stato dell'Arte

L'ultimo decennio è stato caratterizzato da una grande evoluzione metodologica: si è infatti passati da approcci basati su estrazione manuale delle feature e *feature engineering* a sistemi basati su Deep Learning. In questa sezione viene analizzata la letteratura scientifica sulle applicazioni di AI per la diagnosi assistita della malattia di Alzheimer, con l'obiettivo di contestualizzare i capitoli precedenti nel panorama attuale e di sintetizzare i risultati ottenuti dai principali studi di riferimento, le architetture più performanti e, soprattutto, i limiti metodologici ancora presenti come la gestione del Data Leakage.

Fino a pochi anni fa, l'approccio standard richiedeva una pipeline di elaborazione manuale complessa; i metodi di Machine Learning tradizionale, infatti, dipendevano interamente dalla fase di *feature extraction*, in cui bisognava definire le regioni di interesse, come il volume dell'ippocampo o lo spessore corticale, e calcolare i parametri da fornire al classificatore. Sebbene sia un approccio interpretabile, le prestazioni di questo processo dipendono interamente dalla qualità della segmentazione e dalla conoscenza pregressa degli esperti. Con l'avvento del Deep Learning e delle Reti Neurali Convoluzionali, si è assistito ad una svolta. Questo approccio è superiore grazie al cosiddetto *feature learning*: la rete, infatti, apprende autonomamente le caratteristiche importanti direttamente dai voxel dell'immagine, catturando relazioni non lineari che spesso tramite le metriche tradizionali non sono visibili [51].

### 4.1 Rassegna delle applicazioni di AI su dati MRI

L'Imaging a sMRI è la modalità di acquisizione più comune per lo studio dei cambiamenti di morfologia associati alla neurodegenerazione, come per esempio l'atrofia ippocampale [52]. Il passaggio da approcci basati su ROI predefinite a architetture *voxel-based* ha segnato un cambiamento importante. Gli autori Wen et al. nel 2020 hanno condotto un'analisi comparativa delle architetture basate su CNN, da cui è emerso che i modelli che sfruttano input tridimensionali (3D-CNN) tendono a superare le varianti 2D *slice-based* nel catturare caratteristiche spaziali

complesse tipiche della neurodegenerazione [50]. Lo studio ha inoltre stabilito un framework riproducibile per la valutazione, in cui viene sottolineata la criticità del Data Leakage. Gli autori hanno dimostrato, richiamando le criticità metodologiche e i rischi di sovrastima dell'accuratezza discussi nel dettaglio nella Sezione 3.3.2, che separare rigorosamente i dati a livello di soggetto è una prassi necessaria per garantire l'affidabilità clinica dei risultati ottenuti. [50]

Contemporaneamente all'ottimizzazione dei modelli, il miglioramento della robustezza dei classificatori è stato un focus centrale della ricerca, anche attraverso tecniche di *ensemble learning*. Gli autori Islam e Zhang hanno proposto un sistema che prende predizioni di diverse CNN profonde e le aggrega. Questo approccio *ensemble* riduce la varianza rispetto a quella degli errori dei singoli modelli e questo permette di identificare pattern specifici anche nel caso di dataset limitati o sbilanciati [53]. I risultati che hanno riportato dimostrano che combinare estrattori di feature distinti incrementa significativamente le metriche di classificazione nel compito binario "AD vs. CN" rispetto a quelli che si ottengono utilizzando una singola architettura [53].

La sfida più complessa, tuttavia, sta nella diagnosi precoce e nella prognosi dei soggetti affetti da MCI. Uno studio condotto da Basaia et al. ha affrontato questa problematica utilizzando una 3D-CNN applicata direttamente a immagini pesate T1 per discriminare non solo tra AD e CN, ma nello specifico tra MCI stabili (s-MCI) e MCI che convertono in Alzheimer (c-MCI) entro un arco temporale definito [54]. L'importanza di questo studio è l'aver dimostrato che l'approccio DL è in grado di estrarre feature con prestazioni comparabili o superiori ai metodi tradizionali, basati su feature ingegnerizzate manualmente, e l'aver evidenziato come l'analisi automatizzata di tutta la struttura cerebrale fornisca un supporto decisionale valido e capace di identificare i pazienti maggiormente a rischio di conversione in Alzheimer [54].

## 4.2 Rassegna delle applicazioni di AI su dati PET

L'applicazione delle metodologie di AI sulle neuroimmagini ottenute da PET si focalizza su due tipologie di traccianti: il  $^{18}\text{F}$ -FDG per valutare il metabolismo del glucosio e i traccianti specifici per le placche di Amiloide e i grovigli di proteina Tau. L'utilizzo del Deep Learning per l'analisi delle mappe metaboliche si è dimostrato efficiente, con elevate capacità di diagnosi precoce. Nello studio di Ding et al., gli autori hanno adattato l'architettura Inception V3 per poter analizzare scansioni cerebrali 2D. Il modello è stato poi addestrato per classificare lo stato cognitivo (AD, MCI o CN) e predire la progressione verso l'Alzheimer. Un risultato chiave di questo lavoro è il confronto diretto con le prestazioni umane: sensibilità e specificità dell'algoritmo sono risultate superiori a quelle medie di radiologi con differente esperienza; questo ci suggerisce che le reti neurali abbiano la capacità di cogliere pattern di ipometabolismo abbastanza sottili da risultare quasi invisibili all'occhio umano e di conseguenza difficilmente rilevabili all'esame visivo standard [55].

La degenerazione del metabolismo è un processo che presenta pattern spaziali eterogenei e complessi, per questo motivo, per affrontare questa variabilità, Lu et al. hanno proposto un approccio innovativo: gli autori hanno introdotto una rete neurale multiscala (*Multiscale DNN*), progettata per estrarre caratteristiche a diversi livelli di risoluzione. Questa rete permette di integrare informazioni come l'ipometabolismo in specifici voxel con le relazioni presenti tra diverse aree cerebrali, migliorando significativamente l'accuratezza nel differenziare tra soggetti con MCI e controlli sani [56].

Oltre alla valutazione del metabolismo, l'AI trova applicazioni anche nella quantificazione e classificazione dei biomarcatori dell'Alzheimer, come l'accumulo di  $\beta$ -Amiloide. Come evidenziato nella review degli autori Jo, Nho e Saykin, gli algoritmi di Deep Learning vengono sempre più impiegati per cercare di prevedere la "positività all'amiloide", ovvero lo status definito dalla presenza di placche amiloidi, che rappresenta, come ampiamente discusso, uno dei marker biologici fondamentali per la diagnosi di AD. Automatizzare questo processo grazie all'AI permette non solo di standardizzare la classificazione dei pazienti negli studi clinici, ma permette anche di correlare in modo non lineare i pattern di deposizione dell'amiloide con il declino cognitivo [51].

### **4.3 Approcci multimodali: la fusione di dati da strumentazione diversa**

La malattia di Alzheimer è complessa, caratterizzata da atrofia strutturale, contemporaneamente a ipometabolismo e accumulo proteico. Per questo motivo, integrare informazioni provenienti da diverse modalità di imaging potrebbe offrire una capacità discriminativa superiore rispetto ai singoli approcci monomodali. Le architetture di Deep Learning multimodali sfruttano la complementarità delle informazioni MRI (morfologia) e PET (funzionalità), simulando il processo decisionale umano che si basa sull'osservazione e sull'analisi di esami strumentali di diverso tipo.

Un esempio significativo di tale approccio è il lavoro svolto da Huang et al., che nel loro studio propongono un'architettura a due vie. In questo modello, vengono estratte inizialmente le feature di MRI e FDG-PET, separate, da due network *VGG-like*; in seguito, queste vengono concatenate e inserite in diversi layer *fully connected* per la classificazione finale. I risultati che sono stati riportati dimostrano che la fusione delle feature permette al classificatore di trovare una correlazione tra l'atrofia regionale e il corrispondente deficit metabolico, con un'accuratezza che ha raggiunto il 90.10% per la classificazione binaria CN vs. AD, 87.46% per la classificazione CN vs. p-MCI (*progressive MCI*, terminologia equivalente a c-MCI) e 76.90% per il task s-MCI vs. p-MCI, superiori a quelle ottenute utilizzando separatamente dati MRI o PET [57].

La ricerca recente, oltre alla sola fusione di immagini di diversa strumentazione, si sta orientando anche verso l'integrazione di dati eterogenei, unendo il neuroimaging a dati clinici, genetici e demografici. Gli autori Spasov et al. hanno affrontato la sfida della prognosi nell'MCI, utilizzando un approccio che analizza non solo le neuroimmagini, ma integra anche variabili cliniche come età, sesso e genotipo APOE4, direttamente nell'architettura della rete neurale. In questo modo si permette ai modelli di contestualizzare i dati di imaging, associandoli al rischio genetico e demografico del paziente e migliorandone di conseguenza la robustezza nel prevedere la conversione a breve termine (entro 3 anni) [58].

## 4.4 Dataset di riferimento

Per creare e addestrare gli algoritmi di AI servono grandi quantità di dati, che devono essere eterogenei e etichettati con precisione. Nel contesto dell'Alzheimer, si è potuto superare il limite dei dataset ridotti utilizzati in alcuni studi grazie alla creazione di dataset *open-access* standardizzati, così da avere un benchmark comune per confrontare le prestazioni dei modelli.

Il gold standard indiscusso per la ricerca sull'Alzheimer è rappresentato da un dataset già citato nei capitoli precedenti: l'*Alzheimer Disease Neuroimaging Initiative* (ADNI), lanciato per la prima volta nel 2004 [59]. ADNI raccoglie dati clinici, genetici e di neuroimaging di soggetti seguiti nel tempo attraverso visite periodiche. La sua struttura si è evoluta in diverse fasi (ADNI-1, ADNI-GO, ADNI-2, ADNI-3), passando da 800 soggetti nella prima iterazione ADNI-1 a migliaia nella più recente ADNI-3, e include diverse tipologie di dati [59, 60, 61]:

- **Dati di Imaging:** scansioni MRI strutturali (T1-weighted, T2-weighted, FLAIR), acquisite a 1.5 T e 3 T e scansioni PET con traccianti per il glucosio (FDG), per l'Amiloide (AV45/Florbetapir) e per la proteina Tau (AV1451).
- **Biomarcatori non-imaging:** dati genetici, analisi del CSF e test neuropsicologici standardizzati.

La rigorosa standardizzazione dei protocolli di acquisizione è il motivo dell'enorme importanza di ADNI, poiché riduce la variabilità tra i diversi scanner utilizzati e rende i dati adatti per l'addestramento di reti neurali profonde.

In parallelo ad ADNI, esiste anche il progetto OASIS (*Open Access Series of Imaging Studies*), che rappresenta una risorsa importante per la validazione incrociata. La versione più recente offre un dataset di più di 1000 partecipanti. A differenza di ADNI, che è fortemente focalizzato sulla patologia di Alzheimer, OASIS include moltissimi dati riguardo all'invecchiamento normale; questo permette di addestrare modelli che siano capaci di distinguere tra declino cognitivo fisiologico e quello patologico [62]. Utilizzare questi dataset in modo combinato permette

di ridurre il rischio di overfitting su una specifica popolazione o su uno specifico protocollo strumentale, rendendo i modelli più robusti.

## 4.5 Le sfide emergenti

Nonostante i notevoli progressi, integrare i modelli di Deep Learning nella pratica clinica rimane complicato. La letteratura più recente identifica due ostacoli principali che precludono il passaggio dalla ricerca alla pratica ospedaliera: la natura "black box" degli algoritmi e la difficoltà di generalizzazione su dati eterogenei [63].

### 4.5.1 Interpretabilità e Explainable AI

Le reti neurali profonde, in particolare le CNN, sono composte da milioni di parametri e da diverse non-linearità che rendono opaco il processo decisionale del modello, non potendo risalire alle motivazioni della decisione presa. In ambito medico, questa mancanza di trasparenza comporta seri problemi etici e di sicurezza: il radiologo, di fatto, deve poter verificare che la diagnosi del modello si basi su biomarcatori e osservazioni reali e non su artefatti dell'immagine o su allucinazioni. [64].

Per cercare di risolvere questo problema è nato l'ambito della *Explainable AI* (XAI). Le tecniche *post-hoc* di XAI più diffuse in neuroimaging sono per esempio le mappe di salienza, che tentano di evidenziare le regioni dell'input che hanno maggiormente influenzato l'output della rete. Tuttavia, queste spiegazioni possono risultare instabili o poco intuitive per il personale clinico; è quindi necessario sviluppare modelli che siano intrinsecamente interpretabili in applicazioni ad alto rischio e in ambiti estremamente delicati come la diagnosi di Alzheimer [63, 65].

### 4.5.2 Generalizzabilità e Domain Shift

L'applicazione su larga scala è ostacolata da un'ulteriore sfida, che riguarda la robustezza dei modelli quando questi vengono testati su dati provenienti da domini diversi da quelli di addestramento. Per le immagini MRI strutturali, le fonti di variabilità includono [66, 67]:

- **Intensità del campo magnetico:** i dataset clinici spesso combinano acquisizioni a 1.5 T e 3 T. Tuttavia, le immagini a 3 T presentano un SNR superiore e un contrasto diverso rispetto a quelle a 1.5 T. Di conseguenza, un modello addestrato prevalentemente su immagini a campo più elevato potrebbe non riconoscere correttamente le strutture anatomiche in immagini a campo inferiore più rumorose, portando a errori di segmentazione.
- **Differenze hardware e di protocollo:** eventuali variazioni nei parametri di sequenza e nelle caratteristiche costruttive delle bobine di diversi produttori possono causare diffe-

renze nella texture delle immagini. Le reti neurali potrebbero erroneamente interpretare queste differenze come caratteristiche patologiche.

La ricerca attuale, quindi, si sta orientando verso tecniche di *Domain Adaptation* e *Data Augmentation*, che mirano a rendere i modelli più neutrali rispetto allo scanner utilizzato, così da poter garantire le stesse prestazioni indipendentemente dall'hardware.

# Capitolo 5

## Discussione e Conclusioni

### 5.1 Sintesi dei risultati e confronto delle metodologie

L'analisi svolta da questo elaborato ha evidenziato come le applicazioni di AI al neuroimaging dell'Alzheimer si siano profondamente evolute dal punto di vista metodologico, passando da sistemi basati su regole e feature estratte manualmente a modelli automatici di apprendimento. Revisionando la letteratura, è possibile constatare che l'adozione del Deep Learning ha superato i limiti intrinseci del Machine Learning tradizionale, spostando il focus della ricerca dalla selezione manuale delle feature all'apprendimento automatico di pattern.

Un confronto importante è quello tra gli approcci basati sul *Feature Engineering* e quelli basati sul *Feature Learning*. Nonostante studi basati su SVM come quello di Magnin et al. (2009) abbiano avuto grande accuratezza utilizzando feature estratte manualmente da immagini MRI *whole-brain*, questi metodi sono risultati fortemente dipendenti dalla segmentazione delle ROI. Al contrario, le architetture basate su CNN come quelle proposte da Jo, Nho e Saykin (2019) e Basaia et al. (2019) hanno dimostrato di essere capaci di identificare biomarcatori di neurodegenerazione e di positività all'amiloide senza aver bisogno di un intervento umano nella selezione delle feature, ottenendo prestazioni comparabili o superiori e garantendo una maggiore robustezza su dati non segmentati [37, 51, 54].

Invece, per quanto riguarda le architetture di Deep Learning, il confronto tra approcci 2D e 3D sui dati MRI mostra come le CNN 2D abbiano diversi benefici, tra cui un minor costo computazionale – e, di conseguenza, di hardware – e rendano possibile sfruttare il *Transfer Learning*. Tuttavia, lo studio sistematico di Wen et al. (2020) ha mostrato come le CNN 3D siano superiori nel comprendere la struttura tridimensionale dell'atrofia cerebrale. Allo stesso tempo, lo studio mette in guardia sulla validità dei risultati: molte architetture 2D in letteratura con risultati apparentemente eccellenti sono in realtà state affette da Data Leakage, mentre i modelli 3D, se validati con una separazione *subject-based* rigorosa piuttosto che *slice-based*,

offrono una stima delle prestazioni più realistica e affidabile dal punto di vista clinico [50].

Per quanto riguarda l'imaging funzionale con PET, i risultati hanno pienamente confermato che l'AI può superare la capacità di analisi visiva umana. Il modello basato su Inception v3 proposto da Ding et al. (2019) ha avuto sensibilità e specificità superiori a quelle di alcuni radiologi nel task di rilevazione dell'ipometabolismo da FDG-PET. Inoltre, l'approccio multiscala di Lu et al. (2018) ha confermato che integrare informazioni locali (a livello di voxel) e globali (relazioni tra aree cerebrali) migliora significativamente le capacità di distinguere tra controlli sani e soggetti con MCI [55, 56].

Tuttavia, il risultato più importante è stato ottenuto confrontando gli approcci monomodali con quelli multimodali. La malattia di Alzheimer è una patologia che coinvolge molti fattori diversi: atrofia strutturale, alterazioni metaboliche e accumulo di proteine. Lo studio di Huang et al. (2019) ha dimostrato che l'utilizzo combinato delle feature estratte sia da MRI che da FDG-PET porta a un incremento delle prestazioni diagnostiche rispetto all'utilizzo delle due modalità prese singolarmente, con una capacità molto migliorata di distinguere i sottogruppi di p-MCI. Questo viene ulteriormente supportato dal lavoro di Spasov et al. (2019), che ha evidenziato che integrare dati eterogenei - ovvero imaging e dati clinici e genetici - sia importante per la prognosi a breve termine e permetta di prevedere la conversione da MCI a AD con precisione maggiore rispetto ai soli dati di neuroimaging [57, 58].

## 5.2 Criticità e limiti della ricerca attuale

Nonostante, per quanto riguarda l'accuratezza delle diagnosi i risultati siano molto promettenti, integrare i modelli di Deep Learning per la diagnosi di Alzheimer nella pratica clinica rimane complicato a causa di diverse criticità. La letteratura scientifica a riguardo ha mostrato che le prestazioni ottenute non si traducono automaticamente in efficacia clinica, a causa principalmente di quattro fattori: la mancanza di interpretabilità, la scarsa generalizzabilità, i bias metodologici nella validazione e la disponibilità limitata di dati etichettati.

1. **Il problema della "black box" e l'interpretabilità clinica:** La natura di "black box" delle CNN profonde è forse l'ostacolo più grande all'applicazione in campo medico. Come discusso da Rudin (2019), utilizzare modelli "black box", ovvero modelli il cui ragionamento interno è inaccessibile, per decisioni ad alto rischio come una diagnosi di Alzheimer solleva problemi sia etici che di sicurezza, poiché non è possibile verificare il ragionamento logico che sta dietro alla predizione. Sebbene siano state introdotte tecniche di *Explainable AI* come le mappe di salienza, studi come quello di Ghassemi, Oakden-Rayner e Beam (2021) e Singh, Sengupta e Lakshminarayanan (2020) hanno dimostrato

che queste spiegazioni a posteriori possono essere instabili e fornire una "falsa speranza" di comprensione [63, 65, 64].

2. **Generalizzabilità e Domain Shift:** La maggior parte dei modelli in letteratura è addestrata e testata su dataset standardizzati come ADNI, che utilizzano immagini di alta qualità acquisite con protocolli uniformi. La realtà però è diversa e in ambito clinico è presente una grande eterogeneità, sia per la strumentazione che per i parametri di acquisizione utilizzati. Questo fenomeno, il *Domain Shift*, causa un calo delle prestazioni drastico quando un modello viene applicato a dati provenienti da ospedali diversi da quello del training set. Mårtensson et al. (2020) hanno confermato che, in questi casi, l'affidabilità di un modello DL diminuisce significativamente. Questo evidenzia come siano necessarie tecniche di *Domain Adaptation* ancora non sufficientemente sviluppate [66].
3. **Bias metodologici e Data Leakage:** La revisione sistematica condotta da Wen et al. (2020) ha mostrato che gran parte degli studi da loro analizzati ha sofferto di Data Leakage causato da una suddivisione errata dei dataset. Questo ha portato il modello a memorizzare l'anatomia dei soggetti al posto di apprendere le caratteristiche della patologia, così le metriche di accuratezza sono state "gonfiate" artificialmente, ma i risultati non risultano riproducibili su nuovi pazienti [50].
4. **Scarsità e sbilanciamento dei dati:** Infine, c'è anche un problema riguardo la dimensione dei dataset biomedici. A differenza delle milioni di immagini disponibili su dataset generici come ImageNet, infatti, i dataset di neuroimaging per l'AD contano al più poche migliaia di scansioni. Come evidenziato da Tajbakhsh et al. (2016), il rischio di overfitting aumenta se si addestrano reti profonde con milioni di parametri su campioni così ridotti e, sebbene il *Transfer Learning* mitighi parzialmente questo rischio, la scarsità di database etichettati abbastanza vasti rimane un problema [45].

### 5.3 Prospettive future: l'integrazione nella pratica clinica

Per colmare il divario tra i risultati sperimentali e l'effettiva utilità nella pratica clinica, l'evoluzione dei sistemi di supporto decisionale (CAD) dovrà migliorare in tre aspetti fondamentali.

In primo luogo, è necessario superare l'ostacolo della "black box": l'applicazione clinica richiede infatti il passaggio a modelli intrinsecamente interpretabili, abbandonando le spiegazioni a posteriori spesso inaffidabili [65], per fornire al medico ragionamenti logici, causali e verificabili.

In secondo luogo, l'approccio dovrà essere olistico: l'integrazione multimodale di neuroimmagini MRI e PET con l'aggiunta di dati clinici e genetici migliora drasticamente la prognosi, come dimostrato da Huang et al. [57] e Spasov et al. [58] permettendo di emulare il processo deduttivo umano.

Infine, la ricerca dovrà concentrarsi sul garantire la robustezza dei modelli attraverso lo sviluppo di tecniche di Domain Adaptation, per rendere i modelli il più indipendenti possibile dalla variabilità di strumentazione. Questo passo è molto importante, per passare dalla semplice diagnosi alla comprensione del rischio nel MCI, distinguendo i pazienti stabili da quelli che convertiranno in Alzheimer, per poter intervenire tempestivamente con terapie adatte e personalizzabili.

## **5.4 Considerazioni finali**

Concludendo, il lavoro svolto in questa tesi conferma che le applicazioni di Deep Learning alle neuroimmagini non sono più solo un esercizio teorico, ma sono una realtà e una necessità dovuta all'enorme complessità biologica della malattia di Alzheimer. L'analisi della letteratura ha dimostrato che gli algoritmi più moderni possiedono fuori da ogni dubbio la sensibilità necessaria per rilevare pattern morfologici di neurodegenerazione e di alterazioni metaboliche talmente sottili da sfuggire persino all'occhio esperto del radiologo e, pertanto, possono offrire un supporto decisionale laddove la valutazione tradizionale può essere affetta da variabilità inter- e intra-operatore.

Tuttavia, è altrettanto chiaro che la sola accuratezza non è sufficiente affinché l'adozione in ambito clinico diventi realtà. La sfida per l'ingegneria biomedica, ora, non è più rivolta al massimizzare le performance sui dataset standardizzati, ma deve spostarsi e puntare al garantire trasparenza e robustezza nel mondo reale. Il valore aggiunto dell'Intelligenza Artificiale non sta nella sostituzione del radiologo, ma nel supporto a quest'ultimo. La macchina funge da estrattore e decodificatore della grande complessità dei dati multimodali, ma è il medico che governa la decisione clinica finale.

Il passaggio a questo modello decisionale ibrido è un passo fondamentale per trasformare la diagnosi dell'Alzheimer da una semplice procedura per esclusione a un'analisi predittiva accurata, così da offrire al paziente la risorsa più importante di tutte: il tempo per intervenire.

# Bibliografia

- [1] Martina Zvěřová. «Clinical aspects of Alzheimer's disease». en. In: *Clinical Biochemistry* 72 (ott. 2019), pp. 3–6. DOI: 10.1016/j.clinbiochem.2019.04.015.
- [2] Philip Scheltens et al. «Alzheimer's disease». English. In: *The Lancet* 397.10284 (apr. 2021), pp. 1577–1590. DOI: 10.1016/S0140-6736(20)32205-4.
- [3] G. McKhann et al. «Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease». eng. In: *Neurology* 34.7 (lug. 1984), pp. 939–944. DOI: 10.1212/wnl.34.7.939.
- [4] Kenneth M. Langa e Deborah A. Levine. «The Diagnosis and Management of Mild Cognitive Impairment: A Clinical Review». In: *JAMA* 312.23 (dic. 2014), pp. 2551–2561. DOI: 10.1001/jama.2014.13806.
- [5] Marwan N. Sabbagh et al. «Early Detection of Mild Cognitive Impairment (MCI) in Primary Care». en. In: *The Journal of Prevention of Alzheimer's Disease* 7.3 (giu. 2020), pp. 165–170. DOI: 10.14283/jpad.2020.21.
- [6] Leda Volpi et al. «Detecting cognitive impairment at the early stages: The challenge of first line assessment». English. In: *Journal of the Neurological Sciences* 377 (giu. 2017), pp. 12–18. DOI: 10.1016/j.jns.2017.03.034.
- [7] Clifford R. Jack et al. «NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease». In: *Alzheimer's & dementia : the journal of the Alzheimer's Association* 14.4 (apr. 2018), pp. 535–562. DOI: 10.1016/j.jalz.2018.02.018.
- [8] Freddie Márquez e Michael A. Yassa. «Neuroimaging Biomarkers for Alzheimer's Disease». en. In: *Molecular Neurodegeneration* 14.1 (giu. 2019), p. 21. DOI: 10.1186/s13024-019-0325-5.
- [9] Clifford R. Jack et al. «Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup». eng. In: *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 20.8 (ago. 2024), pp. 5143–5169. DOI: 10.1002/alz.13859.
- [10] Tim Hulsen. «Literature analysis of artificial intelligence in biomedicine». In: *Annals of Translational Medicine* 10.23 (dic. 2022), p. 1284. DOI: 10.21037/atm-2022-50.

- [11] Winner Olabiyi, Docas Akinyele e Emmanuel Joel. «The Evolution of AI: From Rule-Based Systems to Data-Driven Intelligence». In: (gen. 2025).
- [12] Peraschi Selvan Subramanian. «The Transformative Role of Artificial Intelligence in Accelerating Biomedical Research: A Focus on Alzheimer’s Disease». en. In: *Journal of Computer Science and Technology Studies* 7.4 (mag. 2025), pp. 375–379. DOI: 10.32996/jcsts.2025.7.4.44.
- [13] Zhaomin Yao et al. «Artificial intelligence-based diagnosis of Alzheimer’s disease with brain MRI images». In: *European Journal of Radiology* 165 (ago. 2023), p. 110934. DOI: 10.1016/j.ejrad.2023.110934.
- [14] Jerrold T. Bushberg et al. *The essential physics of medical imaging*. en. 3rd edition. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2012.
- [15] Giovanni B. Frisoni et al. «The clinical use of structural MRI in Alzheimer disease». In: *Nature reviews. Neurology* 6.2 (feb. 2010), pp. 67–77. DOI: 10.1038/nrneuro1.2009.215.
- [16] Saima Rathore et al. «A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer’s disease and its prodromal stages». In: *NeuroImage* 155 (lug. 2017), pp. 530–548. DOI: 10.1016/j.neuroimage.2017.03.057.
- [17] Carolin Cornelius et al. «Cellular stress response, sirtuins and UCP proteins in Alzheimer disease: role of vitagenes». In: *Immunity & Ageing : I & A* 10 (ott. 2013), p. 41. DOI: 10.1186/1742-4933-10-41.
- [18] Lisa Mosconi. «Brain glucose metabolism in the early and specific diagnosis of Alzheimer’s disease». en. In: *European Journal of Nuclear Medicine and Molecular Imaging* 32.4 (apr. 2005), pp. 486–510. DOI: 10.1007/s00259-005-1762-7.
- [19] Hyun-Kyung Park et al. «Identification of PSEN1 and APP Gene Mutations in Korean Patients with Early-Onset Alzheimer’s Disease». In: *Journal of Korean Medical Science* 23.2 (apr. 2008), pp. 213–217. DOI: 10.3346/jkms.2008.23.2.213.
- [20] L. Harper et al. «An algorithmic approach to structural imaging in dementia». en. In: *Journal of Neurology, Neurosurgery & Psychiatry* 85.6 (giu. 2014), pp. 692–698. DOI: 10.1136/jnnp-2013-306285.
- [21] Jerry L. Prince e Jonathan M. Links. *Medical imaging signals and systems*. en. 2., [updated] ed. Boston: Pearson, 2015.
- [22] Clifford R. Jack et al. «The Alzheimer’s Disease Neuroimaging Initiative (ADNI): MRI methods». eng. In: *Journal of magnetic resonance imaging: JMRI* 27.4 (apr. 2008), pp. 685–691. DOI: 10.1002/jmri.21049.

- [23] Isabella Castiglioni et al. «AI applications to medical images: From machine learning to deep learning». English. In: *Physica Medica: European Journal of Medical Physics* 83 (mar. 2021), pp. 9–24. DOI: 10.1016/j.ejmp.2021.02.006.
- [24] Stephen M. Smith. «Fast robust automated brain extraction». In: *Human Brain Mapping* 17.3 (set. 2002), pp. 143–155. DOI: 10.1002/hbm.10062.
- [25] Boubakeur Belaroussi et al. «Intensity non-uniformity correction in MRI: Existing methods and their validation». In: *Medical Image Analysis* 10.2 (apr. 2006), pp. 234–246. DOI: 10.1016/j.media.2005.09.004.
- [26] Nicholas J. Tustison et al. «N4ITK: Improved N3 Bias Correction». In: *IEEE transactions on medical imaging* 29.6 (giu. 2010), pp. 1310–1320. DOI: 10.1109/TMI.2010.2046908.
- [27] Wang Cong et al. «A Modified Brain MR Image Segmentation and Bias Field Estimation Model Based on Local and Global Information». In: *Computational and Mathematical Methods in Medicine* 2016 (2016), p. 9871529. DOI: 10.1155/2016/9871529.
- [28] John Ashburner e Karl J. Friston. «Voxel-Based Morphometry—The Methods». In: *NeuroImage* 11.6 (giu. 2000), pp. 805–821. DOI: 10.1006/nimg.2000.0582.
- [29] John Ashburner. «SPM: a history». eng. In: *NeuroImage* 62.2 (ago. 2012), pp. 791–800. DOI: 10.1016/j.neuroimage.2011.10.025.
- [30] Mark Jenkinson et al. «FSL». eng. In: *NeuroImage* 62.2 (ago. 2012), pp. 782–790. DOI: 10.1016/j.neuroimage.2011.09.015.
- [31] Bruce Fischl. «FreeSurfer». eng. In: *NeuroImage* 62.2 (ago. 2012), pp. 774–781. DOI: 10.1016/j.neuroimage.2012.01.021.
- [32] Ian Goodfellow, Aaron Courville e Yoshua Bengio. *Deep learning*. eng. Adaptive computation and machine learning. Cambridge, Massachusetts: The MIT Press, 2016.
- [33] Pablo Duboue. *The Art of Feature Engineering: Essentials for Machine Learning*. en. 1<sup>a</sup> ed. Cambridge University Press, mag. 2020. DOI: 10.1017/9781108671682.
- [34] Hema Sekhar Reddy Rajula et al. «Comparison of Conventional Statistical Methods with Machine Learning in Medicine: Diagnosis, Drug Development, and Treatment». eng. In: *Medicina* 56.9 (set. 2020), p. 455. DOI: 10.3390/medicina56090455.
- [35] Muhammad Fahim Uddin et al. «Proposing Enhanced Feature Engineering and a Selection Model for Machine Learning Processes». en. In: *Applied Sciences* 8.4 (apr. 2018), p. 646. DOI: 10.3390/app8040646.
- [36] Corinna Cortes e Vladimir Vapnik. «Support-Vector Networks». en. In: *Machine Learning* 20.3 (set. 1995), pp. 273–297. DOI: 10.1023/A:1022627411411.

- [37] Benoît Magnin et al. «Support vector machine-based classification of Alzheimer’s disease from whole-brain anatomical MRI». en. In: *Neuroradiology* 51.2 (feb. 2009), pp. 73–83. DOI: 10.1007/s00234-008-0463-x.
- [38] Zhen Zhao et al. «Conventional machine learning and deep learning in Alzheimer’s disease diagnosis using neuroimaging: A review». English. In: *Frontiers in Computational Neuroscience* 17 (feb. 2023). DOI: 10.3389/fncom.2023.1038636.
- [39] Yann Lecun et al. «Gradient-Based Learning Applied to Document Recognition». In: *Proceedings of the IEEE* 86 (dic. 1998), pp. 2278–2324. DOI: 10.1109/5.726791.
- [40] A. M. El-Assy et al. «A novel CNN architecture for accurate early detection and classification of Alzheimer’s disease using MRI data». eng. In: *Scientific Reports* 14.1 (feb. 2024), p. 3463. DOI: 10.1038/s41598-024-53733-6.
- [41] Geert Litjens et al. «A survey on deep learning in medical image analysis». eng. In: *Medical Image Analysis* 42 (dic. 2017), pp. 60–88. DOI: 10.1016/j.media.2017.07.005.
- [42] Karen Simonyan e Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. arXiv:1409.1556 [cs]. Apr. 2015. DOI: 10.48550/arXiv.1409.1556.
- [43] Kaiming He et al. «Deep Residual Learning for Image Recognition». en. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, giu. 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.
- [44] Christian Szegedy et al. «Going deeper with convolutions». In: giu. 2015, pp. 1–9. DOI: 10.1109/CVPR.2015.7298594.
- [45] Nima Tajbakhsh et al. «Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?» eng. In: *IEEE transactions on medical imaging* 35.5 (mag. 2016), pp. 1299–1312. DOI: 10.1109/TMI.2016.2535302.
- [46] Jia Deng et al. «ImageNet: a Large-Scale Hierarchical Image Database». In: giu. 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.
- [47] Marina Sokolova e Guy Lapalme. «A systematic analysis of performance measures for classification tasks». In: *Information Processing & Management* 45.4 (lug. 2009), pp. 427–437. DOI: 10.1016/j.ipm.2009.03.002.
- [48] Andrew P. Bradley. «The use of the area under the ROC curve in the evaluation of machine learning algorithms». In: *Pattern Recognition* 30.7 (lug. 1997), pp. 1145–1159. DOI: 10.1016/S0031-3203(96)00142-2.
- [49] Ron Kohavi. «A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection». In: ago. 1995.

- [50] Junhao Wen et al. «Convolutional neural networks for classification of Alzheimer’s disease: Overview and reproducible evaluation». In: *Medical Image Analysis* 63 (lug. 2020), p. 101694. DOI: 10.1016/j.media.2020.101694.
- [51] Taeho Jo, Kwangsik Nho e Andrew J. Saykin. «Deep Learning in Alzheimer’s Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data». English. In: *Frontiers in Aging Neuroscience* 11 (ago. 2019). DOI: 10.3389/fnagi.2019.00220.
- [52] Rosanna Turrisi, Alessandro Verri e Annalisa Barla. «Deep learning-based Alzheimer’s disease detection: reproducibility and the effect of modeling choices». English. In: *Frontiers in Computational Neuroscience* 18 (set. 2024). DOI: 10.3389/fncom.2024.1360095.
- [53] Jyoti Islam e Yanqing Zhang. «Brain MRI analysis for Alzheimer’s disease diagnosis using an ensemble system of deep convolutional neural networks». en. In: *Brain Informatics* 5.2 (mag. 2018), p. 2. DOI: 10.1186/s40708-018-0080-3.
- [54] Silvia Basaia et al. «Automated classification of Alzheimer’s disease and mild cognitive impairment using a single MRI and deep neural networks». en. In: *NeuroImage: Clinical* 21 (2019), p. 101645. DOI: 10.1016/j.nicl.2018.101645.
- [55] Yiming Ding et al. «A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18F-FDG PET of the Brain». eng. In: *Radiology* 290.2 (feb. 2019), pp. 456–464. DOI: 10.1148/radiol.2018180958.
- [56] Donghuan Lu et al. «Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer’s disease». In: *Medical Image Analysis* 46 (mag. 2018), pp. 26–34. DOI: 10.1016/j.media.2018.02.002.
- [57] Yechong Huang et al. «Diagnosis of Alzheimer’s Disease via Multi-Modality 3D Convolutional Neural Network». eng. In: *Frontiers in Neuroscience* 13 (2019), p. 509. DOI: 10.3389/fnins.2019.00509.
- [58] Simeon Spasov et al. «A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer’s disease». In: *NeuroImage* 189 (apr. 2019), pp. 276–287. DOI: 10.1016/j.neuroimage.2019.01.031.
- [59] Susanne Mueller et al. «The Alzheimer’s Disease Neuroimaging Initiative». In: *Neuroimaging clinics of North America* 15 (dic. 2005), pp. 869–77, xi. DOI: 10.1016/j.nic.2005.09.008.
- [60] R. C. Petersen et al. «Alzheimer’s Disease Neuroimaging Initiative (ADNI): clinical characterization». eng. In: *Neurology* 74.3 (gen. 2010), pp. 201–209. DOI: 10.1212/WNL.0b013e3181cb3e25.

- [61] Michael W. Weiner et al. «The Alzheimer’s Disease Neuroimaging Initiative 3: Continued innovation for clinical trial improvement». eng. In: *Alzheimer’s & Dementia: The Journal of the Alzheimer’s Association* 13.5 (mag. 2017), pp. 561–571. DOI: 10.1016/j.jalz.2016.10.006.
- [62] Pamela J. LaMontagne et al. «OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease». en. In: (dic. 2019). Repository: Radiology and Imaging. DOI: 10.1101/2019.12.13.19014902.
- [63] Cynthia Rudin. «Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead». In: *Nature machine intelligence* 1.5 (mag. 2019), pp. 206–215. DOI: 10.1038/s42256-019-0048-x.
- [64] Amitojdeep Singh, Sourya Sengupta e Vasudevan Lakshminarayanan. «Explainable Deep Learning Models in Medical Image Analysis». In: *Journal of Imaging* 6.6 (giu. 2020), p. 52. DOI: 10.3390/jimaging6060052.
- [65] Marzyeh Ghassemi, Luke Oakden-Rayner e Andrew L. Beam. «The false hope of current approaches to explainable artificial intelligence in health care». English. In: *The Lancet Digital Health* 3.11 (nov. 2021), e745–e750. DOI: 10.1016/S2589-7500(21)00208-9.
- [66] Gustav Mårtensson et al. «The reliability of a deep learning model in clinical out-of-distribution MRI data: A multicohort study». In: *Medical Image Analysis* 66 (dic. 2020), p. 101714. DOI: 10.1016/j.media.2020.101714.
- [67] Hao Guan e Mingxia Liu. «Domain Adaptation for Medical Image Analysis: A Survey». In: *IEEE Transactions on Biomedical Engineering* 69.3 (mar. 2022), pp. 1173–1185. DOI: 10.1109/TBME.2021.3117407.