



**Politecnico  
di Torino**

**Politecnico di Torino**

Ingegneria Biomedica  
A.a. 2024/2025  
Sessione di Laurea luglio 2025

# **Diagnosi tra adenoma e carcinoma surrenalico tramite metodi di intelligenza artificiale**

Relatore:

Salvi Massimo

Candidati:

Cannata Marco

Matricola 308676



# ABSTRACT

---

Negli ultimi anni, l'intelligenza artificiale ha compiuto progressi significativi in diversi ambiti, tra cui la medicina, dove viene sempre più impiegata per supportare i medici nella diagnosi, nelle decisioni terapeutiche e nell'ottimizzazione dei processi clinici. In particolare, l'analisi di immagini mediche tramite reti neurali ha rivoluzionato il campo della diagnostica, consentendo di automatizzare la classificazione di tessuti, identificare precocemente patologie e fornire un supporto oggettivo alla valutazione clinica.

Il presente lavoro si propone di contribuire a questa evoluzione sviluppando un modello di deep learning basato su una rete neurale convoluzionale, nello specifico una ResNet-18, per la classificazione e l'analisi di immagini istologiche del tessuto surrenale colorate con una tecnica speciale che evidenzia la struttura della reticolina. L'obiettivo è fornire agli operatori sanitari uno strumento di supporto per migliorare l'identificazione e la caratterizzazione delle masse tumorali del surrene.

I tumori surrenalici rappresentano una condizione relativamente comune nella popolazione mondiale e, sebbene nella maggior parte dei casi siano di natura benigna, in alcune situazioni possono evolvere in forme maligne. La difficoltà principale nella gestione di questi tumori risiede proprio nella determinazione della loro malignità, un aspetto cruciale per definire il percorso terapeutico più adeguato. Per affrontare questa sfida, il modello proposto non si limiterà a classificare le immagini, ma genererà mappe di calore che evidenzieranno visivamente le aree di interesse, distinguendo i diversi tipi di tessuto e facilitando l'interpretazione da parte degli specialisti.

Il lavoro sarà articolato in diverse sezioni. Nei capitoli iniziali verrà fornita una panoramica sui tumori del surrene, illustrandone l'epidemiologia, le caratteristiche cliniche e i principali metodi diagnostici, con un focus particolare sulla colorazione reticolare e sul suo ruolo nell'analisi istopatologica. Successivamente, verrà descritta in dettaglio la preparazione del dataset, le strategie di preprocessing delle immagini e le tecniche di data augmentation adottate per migliorare la generalizzazione del modello. Seguirà la spiegazione della rete neurale utilizzata, approfondendo la struttura della ResNet-18, le modifiche implementate per ottimizzare le prestazioni e i criteri adottati per l'addestramento. Infine, verranno presentati e discussi i risultati ottenuti, analizzando le prestazioni del modello, i limiti riscontrati e le possibili direzioni future per migliorare ulteriormente l'approccio proposto.

Grazie all'integrazione dell'intelligenza artificiale nell'analisi istologica, questo lavoro si propone di offrire un contributo concreto alla diagnosi precoce e alla comprensione delle neoplasie surrenaliche, con l'obiettivo di supportare il personale medico nell'identificazione di tumori maligni e nel miglioramento dell'accuratezza diagnostica.



# INDICE

---

Abstract .....	3
Indice delle figure .....	7
Indice delle Tabelle .....	9
1 Introduzione .....	11
1.1 Ghiandole surrenali .....	11
1.2 Tumore al surrene .....	13
1.2.1 Algoritmo reticolina .....	15
1.3 Digital Pathology .....	17
2 Dataset .....	19
2.1 Patch extraction .....	21
2.2 Data Augmentation & Reduction .....	23
3 Training .....	24
3.1 CNN - Struttura e Funzionamento .....	24
3.2 Resnet .....	26
3.2.1 Skip Connection .....	26
3.2.2 Architettura ResNet18 .....	27
3.3 Parametri di Training .....	33
3.3.1 Dropout .....	33
3.3.2 Regolarizzazione <b>L2</b> .....	34
3.3.3 Ottimizzatore - SGD .....	35
3.3.4 Loss Function - Cross Entropy Loss .....	36
3.4 Criteri di Valutazione .....	38
4 Heatmap .....	40
4.1 Creazione Heatmap .....	40
4.2 Criteri di Valutazione delle Heatmap .....	44

5	Risultati.....	45
5.1	Risultati Allenamento .....	45
5.2	Risultati Heatmap.....	47
6	Conclusioni.....	50
7	Riferimenti.....	51
	Ringraziamenti.....	55

# INDICE DELLE FIGURE

---

Figure 1: Ghiandole surrenali.-----	11
Figure 2: Struttura surrenale e produzione endocrina. -----	12
Figure 3: Schema riassuntivo dell'iter utile a stabilire la malignità e l'attività biochimica della massa surrenale trovata incidentalmente [11]. -----	14
Figure 4: Esempio di estrazione di patch, a sinistra è riportata la label che guida il processo, a destra le patch corrispondenti sulla Tile.-----	21
Figure 5: Esempio di Skip connection.-----	26
Figure 6: Struttura tipica di una ResNet18.-----	27
Figure 7: Architettura dettagliata della ResNet18 utilizzata con nomi dei layer e caratteristiche degli stessi. (Parte I) -----	28
Figure 8: Architettura dettagliata della ResNet18 utilizzata con nomi dei layer e caratteristiche degli stessi. (Parte II) -----	29
Figure 9: Architettura dettagliata della ResNet18 utilizzata con nomi dei layer e caratteristiche degli stessi. (Parte III) -----	30
Figure 10: Architettura dettagliata della ResNet18 utilizzata con nomi dei layer e caratteristiche degli stessi. (Parte IV) -----	31
Figure 11: Architettura dettagliata della ResNet18 utilizzata con nomi dei layer e caratteristiche degli stessi. (Parte V) -----	32
Figure 12: Esempio grafico del dropout e dell'eliminazione dei neuroni e delle loro connessioni durante la fase di training. -----	34
Figure 13: Esempio di Confusion Matrix. -----	38
Figure 14: Visualizzazione della tecnica del Padding.-----	40
Figure 15: Visualizzazione della Patch Extraction. -----	41
Figure 16: Visualizzazione della sovrapposizione delle patch e della zona di intersezione soggetta a classificazione per majority voting. -----	42
Figure 17: Confronto tra la Maschera in scala di grigi (a) e la Heatmap (b) ottenuta con l'applicazione della colormap 'JET'. -----	42
Figure 18: Valori della funzione di perdita sul validation set. -----	45
Figure 19: Risultato parziale di classificazione di una Tile NAM; da sinistra a destra: Tile originale, label manuale e Maschera in scala di grigi. -----	47

Figure 20: Risultato parziale di classificazione di una Tile NAC; da sinistra a destra: Tile originale, label manuale e Maschera in scala di grigi. -----	47
Figure 21: Risultato parziale di classificazione di una Tile ACC, da sinistra a destra: Tile originale, label manuale e Maschera in scala di grigi. -----	48
Figure 22: Risultato parziale di classificazione di una Tile ACA; da sinistra a destra: Tile originale, label manuale e Maschera in scala di grigi. -----	48
Figure 23: Heatmap della Tile NAM. -----	49
Figure 24: Heatmap della Tile NAC. -----	49
Figure 25: Heatmap della Tile ACC. -----	49
Figure 26: Heatmap della Tile ACA. -----	49

# INDICE DELLE TABELLE

---

Tabella 1: Confronto riassuntivo tra ACA e ACC. -----	15
Tabella 2: Dataset originale. -----	19
Tabella 3: Classi originali del dataset, immagini a sinistra e label a destra. -----	20
Tabella 4: Dimensioni del dataset in seguito alla patch extraction e prima di data augmentation e data reduction. -----	22
Tabella 5: Struttura finale del dataset utilizzato per l'addestramento della rete. -----	23
Tabella 6: Pesi delle classi applicate alla Weighted Cross-Entropy Loss Function. -----	37
Tabella 7: Confusion Matrix sul test set. -----	46
Tabella 8: Valori di precision e recall per classe. -----	46

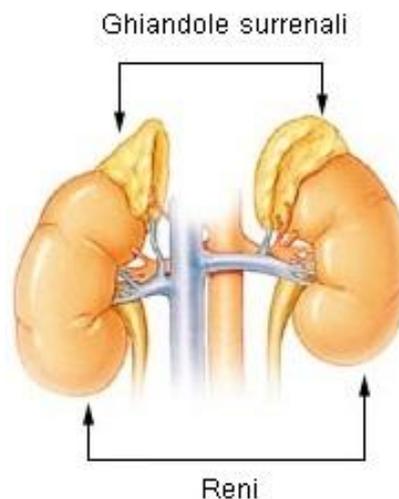


# 1 INTRODUZIONE

---

## 1.1 GHIANDOLE SURRENALI

Le ghiandole surrenali, o surreni, sono due piccoli organi endocrini di forma piramidale situati sopra il polo superiore di ciascun rene. Nonostante le loro ridotte dimensioni (circa 4–5 cm di lunghezza), esse svolgono funzioni endocrine fondamentali per il mantenimento dell'omeostasi, influenzando il metabolismo, la risposta allo stress e l'equilibrio idro-elettrolitico [1] [2].



*Figure 1: Ghiandole surrenali.*

Ogni surrene è composto da due porzioni morfologicamente e funzionalmente distinte: la corteccia surrenale (circa il 90% del volume) e la midollare surrenale, derivanti rispettivamente dal mesoderma e dalla cresta neurale [3] [4].

- **Corteccia surrenale:** è suddivisa in tre zone concentriche che producono ormoni steroidei diversi:
  - La zona glomerulare, la più esterna, è responsabile della secrezione di aldosterone, un mineralcorticoide che regola il riassorbimento di sodio e la secrezione di potassio a livello renale, influenzando direttamente il volume plasmatico e la pressione arteriosa [3] [4].
  - La zona fascicolata, intermedia e la più estesa, sintetizza principalmente cortisolo, un glucocorticoide implicato nella regolazione del metabolismo glucidico, lipidico e

proteico, oltre che nella modulazione della risposta immunitaria e infiammatoria [2] [4].

- o La zona reticolare, più interna, produce androgeni surrenalici come il deidroepiandrosterone (DHEA), che costituiscono una fonte secondaria di ormoni sessuali e possono essere convertiti in testosterone o estrogeni nei tessuti periferici [3] [4].

- **Midollare surrenale:** costituita da cellule cromaffini, secerne catecolamine (principalmente adrenalina e noradrenalina) in risposta a stimoli simpatici. Questi ormoni sono coinvolti nella risposta "fight or flight", aumentando la frequenza cardiaca, la gittata cardiaca, la pressione arteriosa e la disponibilità di glucosio nel sangue [5].

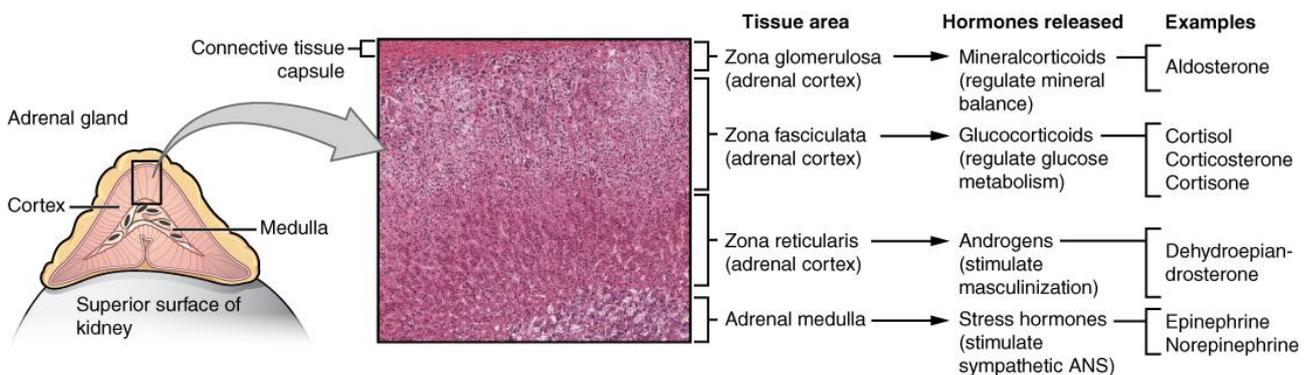


Figure 2: Struttura surrenale e produzione endocrina.

Sebbene le piccole dimensioni, il surrene è soggetto a numerose disfunzioni tipicamente causate da iperfunzione o ipofunzione della ghiandola stessa, come: Morbo di Addison, insufficienza surrenalica secondaria, Sindrome di Cushing, iperaldosteronismo primario (Sindrome di Conn), ecc., o patologie strutturali come le neoplasie.

## 1.2 TUMORE AL SURRENE

Il tumore al surrene rappresenta una condizione patologica di rilevante interesse medico, con un'incidenza che varia in funzione dell'età. In età adulta, la sua prevalenza si attesta intorno al 3% della popolazione mondiale, mentre negli anziani può raggiungere anche il 10% [6]. La maggior parte di questi tumori si manifesta sotto forma di adenomi adrenocorticali (ACA), lesioni di piccole dimensioni, caratterizzate da una crescita lenta e dall'assenza di caratteristiche invasive; nella maggior parte dei casi, risultano non funzionali [7]. Tuttavia, vi sono anche forme più aggressive che presentano un'elevata capacità di proliferazione, infiltrazione locale e metastatizzazione, ovvero i carcinomi adrenocorticali (ACC), estremamente rari e con un'incidenza compresa tra 0.7 – 2 casi per milione all'anno [8].

Sebbene gli ACC siano tipicamente associati all'età adulta, studi epidemiologici hanno evidenziato un secondo picco di incidenza in età pediatrica. Infatti, i carcinomi adrenocorticali rappresentano circa l'1.3 % di tutti i tumori diagnosticati nei bambini, un dato che si contrappone alle statistiche relative all'incidenza di questa neoplasia nella popolazione adulta [9].

La scoperta di un tumore al surrene avviene nella maggior parte dei casi in modo fortuito, spesso durante esami di imaging trasversale come la Tomografia Computerizzata (TC) o la Risonanza Magnetica (RM), eseguita per indagare altre malattie. Proprio per questa ragione, le lesioni surrenaliche asintomatiche di nuova diagnosi vengono definite come '*Adrenal Incidentaloma*' [7] [10]. Una volta identificata una massa surrenale, è fondamentale procedere con una valutazione accurata delle sue caratteristiche, per determinare la funzionalità e, soprattutto, la possibile natura maligna.

Inizialmente, viene eseguita una TC senza mezzo di contrasto, che permette di valutare dimensione, densità e omogeneità della massa; nel caso in cui le caratteristiche della lesione risultino sospette, possono essere necessari ulteriori controlli [11].

Dal punto di vista endocrinologico, circa il 60 % dei tumori surrenalici si presenta come funzionale, ovvero in grado di secernere ormoni. Tra le sostanze coinvolte nella secrezione incontrollata si annoverano aldosterone, ormoni sessuali e i loro precursori, nonché il cortisolo, la cui iperproduzione può determinare l'ipercortisolismo [12]. Questa condizione endocrina è l'alterazione più frequente correlata agli ACC ed è spesso associata a una prognosi sfavorevole. Per tale motivo,

oltre alla morfologia, è necessario indagare riguardo la funzionalità endocrina della massa, verificando l'eventuale iperproduzione ormonale attraverso esami biochimici specifici.

Nel caso in cui la patologia venga diagnosticata in uno stadio ancora localizzato, il trattamento di elezione prevede la resezione chirurgica radicale del tumore. Tuttavia, nonostante l'intervento, la prognosi rimane incerta, poiché la maggior parte dei pazienti (circa il 70-80 %) manifesta recidive entro i primi cinque anni successivi all'operazione [12].

Le stime di sopravvivenza variano in funzione di molteplici fattori, tra cui lo stadio della malattia al momento della diagnosi, la risposta al trattamento e la presenza di fattori di rischio aggiuntivi [7] [12]. Un'accurata valutazione clinica e il monitoraggio costante dei pazienti risultano dunque essenziali per ottimizzare la gestione terapeutica e migliorare le prospettive di sopravvivenza.

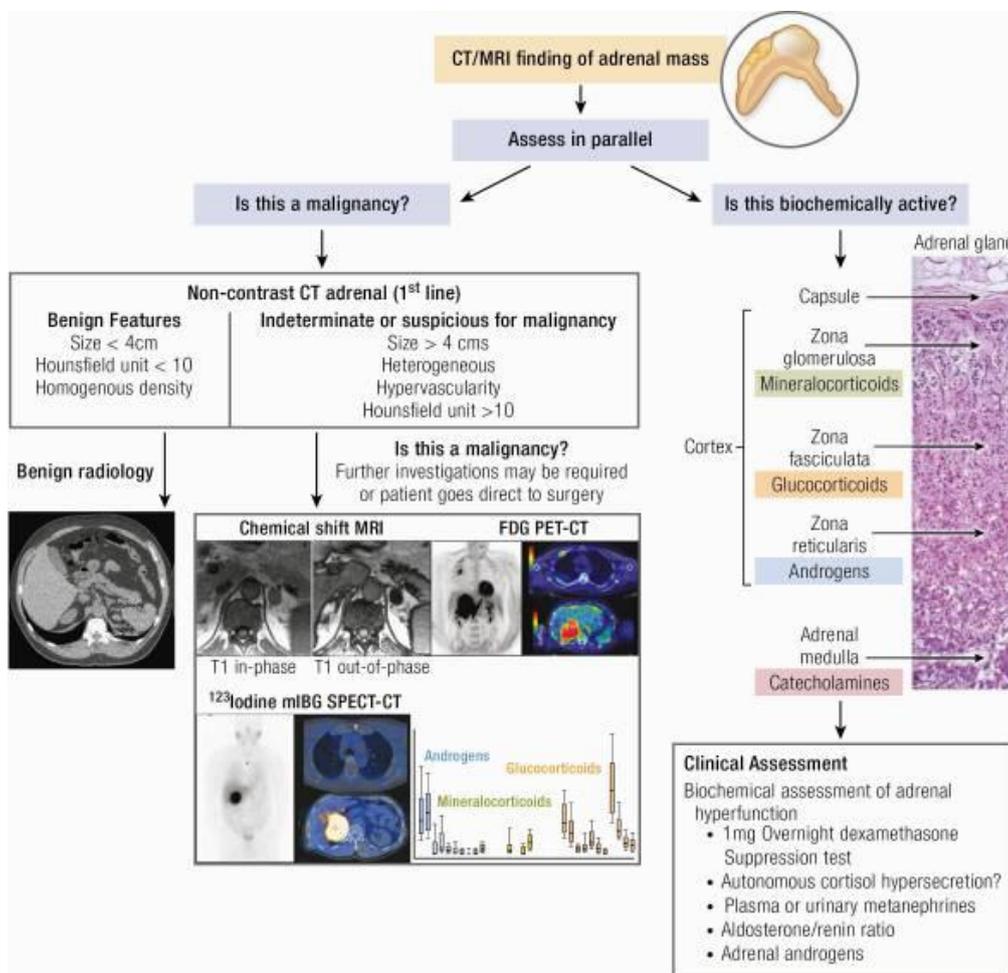


Figure 3: Schema riassuntivo dell'iter utile a stabilire la malignità e l'attività biochimica della massa surrenale trovata incidentalmente [11].

## 1.2.1 ALGORITMO RETICOLINA

I tumori surrenalici, come già accennato, si distinguono tra gli adenomi adrenocorticali e i carcinomi adrenocorticali; la distinzione tra queste due forme è cruciale per la gestione clinica del paziente, poiché un carcinoma surrenalico richiede trattamenti più aggressivi, mentre un adenoma può essere monitorato senza necessità di interventi immediati. Nella tabella sottostante (*Tabella 1*), per una visione più completa vengono riportate le principali differenze tra i due casi:

Caratteristica	ACA	ACC
<b>Natura</b>	Tumore benigno	Tumore maligno e aggressivo
<b>Incidenza</b>	Comune (fino al 10% della popolazione adulta)	Raro (1-2 casi per milione all'anno)
<b>Crescita</b>	Lenta, senza invasione	Rapida, con invasione locale e metastasi
<b>Dimensioni</b>	Solitamente <4 cm	Spesso >6 cm
<b>Produzione ormonale</b>	Può essere non secernente o secernente	Spesso secernente con produzione incontrollata di ormoni
<b>Diagnosi</b>	TAC/RM: bassa densità lipidica, margini regolari	TAC/RM: irregolare, invasivo, alto indice mitotico
<b>Criteri diagnostici</b>	Punteggio di Weiss <3 (benigno)	Punteggio di Weiss $\geq$ 3 (maligno)
<b>Trattamento</b>	Monitoraggio o chirurgia se sintomatico	Surrenectomia, mitotano, chemioterapia
<b>Prognosi</b>	Ottima, raramente recidiva	Sopravvivenza a 5 anni: 20-40%, dipende dallo stadio

*Tabella 1: Confronto riassuntivo tra ACA e ACC.*

Negli anni sono stati sviluppati diversi metodi diagnostici per discriminare tra ACA e ACC con la maggior accuratezza possibile. In *Figura 3* sono riportati alcuni metodi morfologici, ma negli ultimi anni i metodi istologici hanno assunto un'importanza sempre maggiore.

Tra i metodi più utilizzati vi è il punteggio di Weiss; un sistema che si fonda sull'analisi microscopica di nove parametri morfologici specifici. Secondo questo criterio, un tumore viene classificato come carcinoma e quindi considerato maligno se almeno tre di questi parametri sono presenti. In caso contrario, la neoplasia viene definita adenoma e ritenuta benigna.

Nonostante, l'affidabilità del sistema di Weiss, esistono alcune criticità legate alla soggettività nella valutazione di alcuni parametri, soprattutto nei casi limite in cui vengono identificati solo uno o due parametri morfologici. In questi casi sebbene il tumore venga diagnosticato come benigno, esiste un potenziale rischio di malignità residua [10]. Inoltre, la riproducibilità del punteggio di Weiss si riduce in alcune varianti morfologiche del ACC, come la variante oncocitica e la variante mixoide. È stato introdotto il criterio di Weiss modificato per le neoplasie oncocitiche che tiene conto delle peculiarità morfologiche della variante, ma nel caso della variante mixoide, alcuni parametri risultano difficili da identificare, rendendo necessario l'impiego di altre metodologie diagnostiche [13].

Negli ultimi anni, è stato proposto un nuovo metodo di valutazione noto come *Algoritmo della Reticolina* (RA) [13], che si basa sull'osservazione della struttura della reticolina. Questo approccio fornisce un importante parametro qualitativo per distinguere gli ACC dagli ACA, poiché nei carcinomi surrenalici si osserva una distruzione significativa della struttura della reticolina, mentre negli adenomi essa rimane intatta o presenta alterazioni minime.

Il metodo RA si sviluppa in due fasi:

- I. **Colorazione con argento:** inizialmente, il tessuto viene sottoposto a una colorazione speciale a base di argento, che permette di evidenziare la disposizione e l'integrità delle fibre di reticolina.
- II. **Valutazione quantitativa:** se viene rilevata una distruzione della struttura reticolinica, si procede con un'analisi più approfondita, che include tre criteri quantitativi correlati alla malignità: presenza di necrosi, alta frequenza mitotica e invasione venosa.

Uno dei principali vantaggi dell'algoritmo della reticolina è la sua elevata riproducibilità inter-osservatore, rendendolo un metodo diagnostico affidabile e meno soggetto a interpretazioni soggettive rispetto al punteggio di Weiss. Inoltre, la sua rapidità di esecuzione e la facilità di interpretazione ne fanno un valido strumento complementare per la diagnosi differenziale tra ACC e ACA.

## 1.3 DIGITAL PATHOLOGY

Negli ultimi anni, l'intelligenza artificiale ha assunto un ruolo sempre più rilevante nel campo della diagnostica medica, compresa la diagnosi delle patologie legate al surrene. Le sue applicazioni stanno dimostrando un impatto significativo soprattutto nell'analisi delle immagini radiologiche, dove gli algoritmi di deep learning permettono di supportare i medici nell'identificazione accurata di diverse condizioni cliniche.

Tra gli esempi più significativi:

- **Segmentazione automatica delle ghiandole surrenali:** un modello basato su deep learning è stato sviluppato per segmentare in maniera autonoma le ghiandole surrenali nelle scansioni CT, rivelandosi particolarmente efficace nell'identificare casi di iperplasia [14].
- **Riconoscimento degli adenomi surrenalici:** reti neurali convoluzionali profonde (DCNN) sono state impiegate per identificare con successo gli adenomi surrenalici nelle immagini CT, aprendo la strada a diagnosi differenziali ad alta precisione [15].
- **Classificazione di masse surrenali indeterminate:** modelli di deep learning sono stati impiegati per distinguere tra carcinomi adrenocorticali e adenomi poveri di lipidi, affrontando una delle sfide più complesse nella diagnosi radiologica [16].
- **Caratterizzazione dei tumori surrenali funzionali:** l'IA è stata utilizzata per analizzare immagini CT con mezzo di contrasto, riuscendo a quantificare e distinguere i principali tipi di neoplasie endocrine in modo accurato e automatizzato [17].
- **Analisi texture in CT non contrastata:** tecniche di machine learning applicate all'analisi texture hanno permesso di differenziare feocromocitomi subclinici da adenomi poveri di lipidi, anche in assenza di mezzo di contrasto [18].

Negli ultimi anni, anche la classificazione automatica delle immagini istologiche ha attirato crescente attenzione grazie all'avvento di modelli di deep learning sempre più accurati. Diverse architetture CNN, come ResNet, Inception e DenseNet, sono state impiegate per distinguere tessuti normali da quelli tumorali in immagini provenienti da biopsie [19] [20] [21]. Tuttavia, gran parte di questi studi si concentra su tumori comuni come il carcinoma polmonare, al seno o del colon, mentre neoplasie rare come i tumori del surrene rimangono poco esplorate. Inoltre, alcuni approcci richiedono dataset molto grandi o fanno uso di pesi pre-addestrati, rendendo difficile l'adattamento a contesti clinici specifici. Alcuni studi, ad esempio, fanno uso di reti generative avversarie condizionali (cGAN) per generare immagini istologiche sintetiche, contribuendo a migliorare

l'interpretabilità dei modelli di deep learning nella classificazione di vari tumori [22]. Nonostante ciò, la letteratura sull'uso dell'IA nell'istopatologia delle patologie surrenali risulta ancora limitata.

## 2 DATASET

---

Sono state raccolte sedici Wide Slide Image (WSI) di tessuto surrenalico che evidenziano la struttura della reticolina tramite la colorazione speciale specifica per l'Algoritmo della Reticolina.

Le immagini sono così suddivise:

- 7 WSI ACA
- 5 WSI ACC
- 4 WSI di surrenale normale

Da queste WSI sono state estratte 1371 Tiles con dimensioni 1000x1000, con un ingrandimento 20x. Tutte le Tiles estratte contengono tessuto surrenalico con colorazione reticolare per evidenziare la struttura della reticolina. Successivamente, le immagini sono state classificate manualmente in quattro classi mediante valori di grigio, mentre le zone nere rappresentano altri tessuti (*Tabella 3*). Le label, ovvero immagini con classificazione in scala di grigi, sono state suddivise in: Training, Validation e Test set.

	Numero Tiles
Training Set	916
Validation Set	218
Test Set	237
Totale Immagini	1371

*Tabella 2: Dataset originale.*

Tuttavia, il dataset così formato non risulta idoneo all'addestramento di una rete neurale per due motivi principali:

- 1) Le dimensioni delle immagini (1000x1000) sono eccessive, rendendo difficile l'individuazione di particolari utili all'apprendimento e alla classificazione.
- 2) Il dataset presenta uno sbilanciamento pronunciato tra le classi, penalizzando le prestazioni della rete neurale.

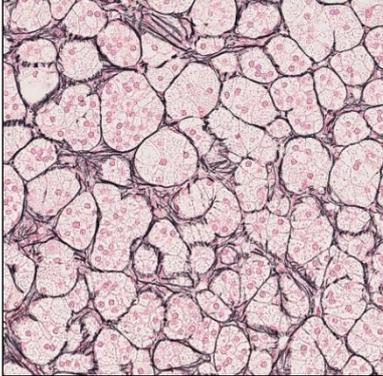
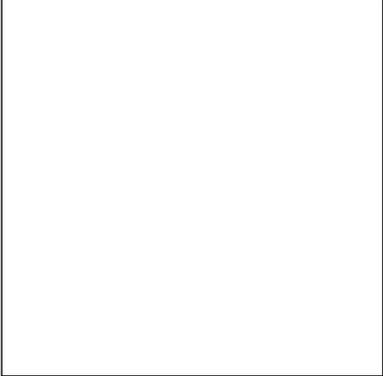
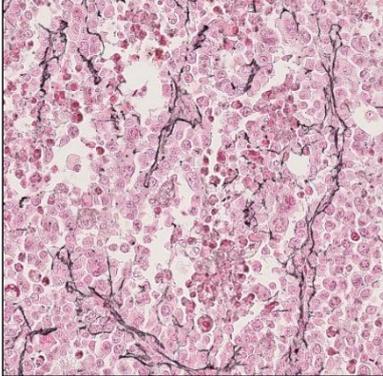
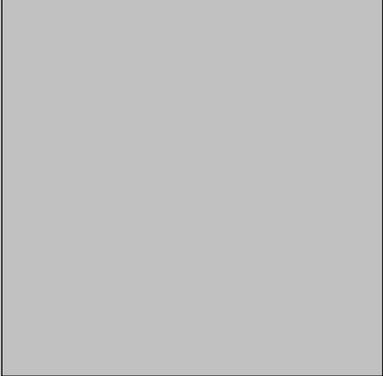
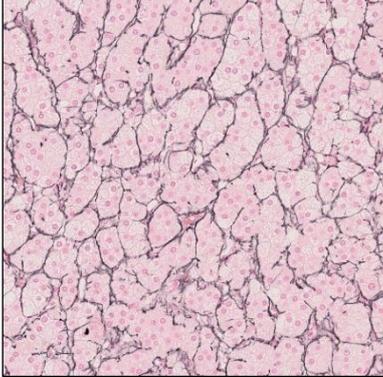
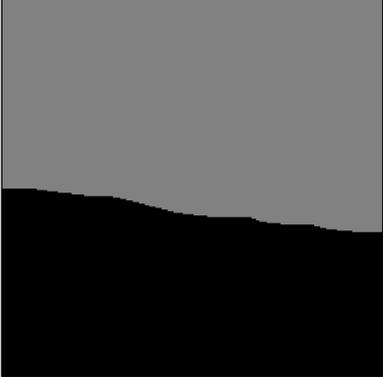
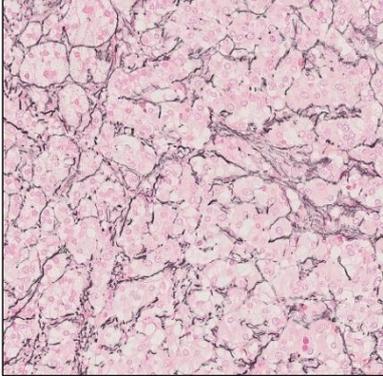
Classe	Valore di grigio
 <p data-bbox="268 656 716 689">ACA (Adrenal Cortical Adenoma)</p>	 <p data-bbox="1078 656 1134 689">255</p>
 <p data-bbox="261 1088 724 1122">ACC (Adrenal Cortical Carcinoma)</p>	 <p data-bbox="1078 1088 1134 1122">192</p>
 <p data-bbox="288 1520 695 1554">NAC (Normal Adrenal Cortex)</p>	 <p data-bbox="1078 1520 1134 1554">128</p>
 <p data-bbox="276 1953 708 1986">NAM (Normal Adrenal Medulla)</p>	 <p data-bbox="1078 1953 1126 1986">64</p>

Tabella 3: Classi originali del dataset, immagini a sinistra e label a destra.

## 2.1 PATCH EXTRACTION

Per affrontare il primo problema, si è scelto di effettuare una *'label-based patch extraction'*, una tecnica di estrazione e divisione intelligente delle immagini; in questo caso eseguita in funzione della classificazione presente nella label corrispondente. La dimensione delle immagini da estrarre è stata selezionata pari a 384x384 pixel con una sovrapposizione delle stesse del 20%.

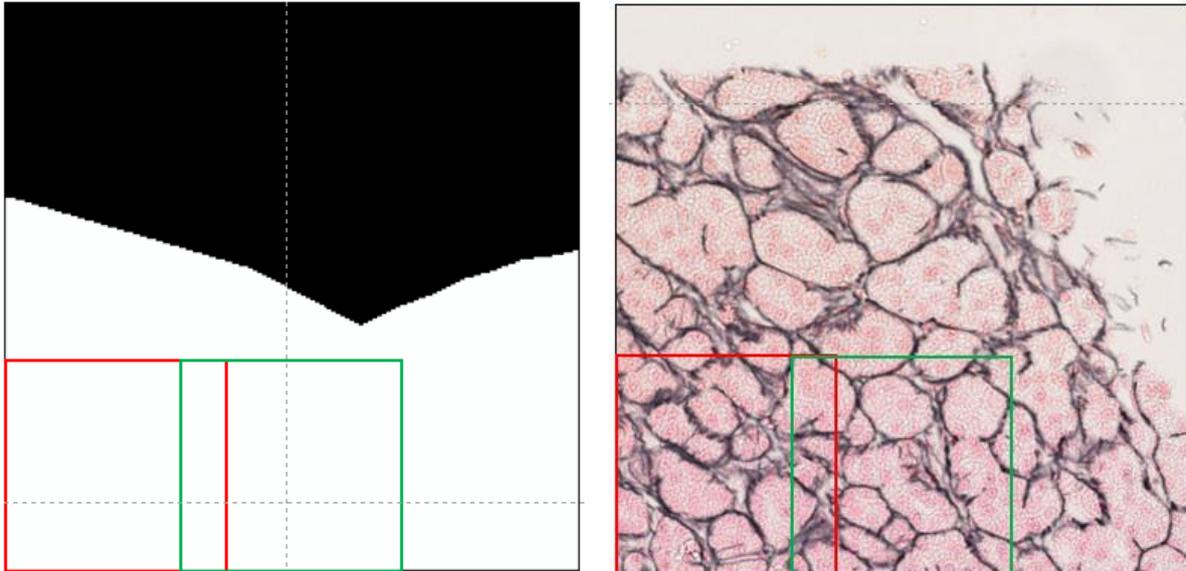


Figure 4: Esempio di estrazione di patch, a sinistra è riportata la label che guida il processo, a destra le patch corrispondenti sulla Tile.

Questa scelta ha permesso di:

- Avere un numero contenuto di patch, evitando di renderlo eccessivamente grande e complesso in vista dell'addestramento della rete e dell'hardware utilizzato.
- Avere un buon compromesso fra dimensioni e visibilità dei dettagli.
- Sottoporre più volte all'addestramento della rete alcune porzioni delle immagini per migliorarne la capacità predittiva.

Sono state sottoposte a questo processo tutte le Tile del dataset e ognuna di esse è stata divisa in un numero variabile di patch, fino ad un massimo di nove<sup>1</sup>. Inoltre, con questa suddivisione è stata

---

<sup>1</sup> Mille pixel dell'immagine originale per lato possono essere divisi, considerando la sovrapposizione applicata, in massimo tre patch della dimensione scelta, lo stesso vale per l'altro lato.

aggiunta una nuova classe, 'other\_tissue (OT)', comprendente tutti i tessuti non classificati nelle categorie riportate nella *Tabella 3* e rappresentati con colorazione nera nelle label.

Grazie alla tecnica utilizzata, ogni patch estratta avrà classificazione univoca in una delle cinque classi; sfruttando questa caratteristica è stato possibile riorganizzare il dataset, mantenendo la struttura dei subset Training, Validation e Test set, internamente organizzati con un numero di cartelle pari al numero delle classi.

La quantità di immagini ottenute in seguito alla patch extraction, suddivise per classi, è riportata nella seguente tabella:

	OT	NAM	NAC	ACC	ACA	Tot.
Training Set	90	47	467	2562	4277	7443
Validation Set	23	13	55	862	855	1808
Test Set	12	17	118	875	859	1881
Totale Immagini						11132

*Tabella 4: Dimensioni del dataset in seguito alla patch extraction e prima di data augmentation e data reduction.*

## 2.2 DATA AUGMENTATION & REDUCTION

Dall'analisi della distribuzione delle immagini, si evince che il dataset non è ancora bilanciato in modo ottimale. In particolare,

- Nel Training set, la classe ACA è sovrarappresentata rispetto alla classe ACC, mentre le altre tre classi, NAC, NAM e OT, sono numericamente inferiori.
- Nel Validation set, invece, le classi ACC e ACA risultano numericamente bilanciate, ma, come per il training set, le altre classi non sono adeguatamente rappresentate.

Per porre rimedio a tale discrepanza numerica e poter migliorare la fase successiva di addestramento della rete sono state applicate due tecniche: Data Augmentation e Data Reduction; la prima, utilizzata per aumentare il numero delle immagini delle classi meno rappresentate nel dataset (OT, NAM e NAC), mentre la seconda, è stata utilizzata per ridurre la numerosità della classe ACA nel Training set.

Le trasformazioni geometriche utilizzate per aumentare la quantità di dati sono:

- Rotazione di 90°,
- Rotazione di 270°,
- Mirroring o riflessione orizzontale,
- Flipping o riflessione verticale.

Queste operazioni hanno permesso di generare quattro nuove immagini per ciascuna immagine originale, permettendo di arricchire il dataset.

La data Reduction, invece, è stata effettuata andando a rimuovere<sup>2</sup> in maniera casuale dal dataset 1677 immagini ACA. Dopo le modifiche apportate, la nuova distribuzione del dataset è la seguente:

	OT	NAM	NAC	ACC	ACA	Tot.
Training Set	450	235	2335	2562	2600	8182
Validation Set	115	65	275	862	855	2172
Test Set <sup>3</sup>	12	17	118	875	859	1881
Totale Immagini						12235

Tabella 5: Struttura finale del dataset utilizzato per l'addestramento della rete.

<sup>2</sup> Le immagini non sono state eliminate del tutto, ma semplicemente allocate in una cartella diversa.

<sup>3</sup> Il test set non è stato sottoposto a queste tecniche perché la distribuzione delle immagini non influisce con il suo ruolo.

# 3 TRAINING

---

Definito il dataset, si può procedere alla parte più lunga e importante del lavoro ovvero l'allenamento di una rete neurale convoluzionale, in inglese Convolutional Neural Network (CNN). Le CNN sono reti ispirate ai processi biologici [23]; a differenza delle reti tradizionali permettono di analizzare ed elaborare con una buona efficacia dati strutturati, come le immagini, estrapolando pattern complessi. Questa capacità le rende sempre più utilizzate in campo medico e in particolare nell'interpretazione delle immagini derivanti da imaging medico.

## 3.1 CNN - STRUTTURA E FUNZIONAMENTO

Le CNN sono formate da più strati sequenziali e gerarchicamente ordinati con lo scopo di processare e trasformare i dati in input tramite l'estrazione delle caratteristiche (features) da un livello base a uno più complesso. Tipicamente, gli strati presenti in una CNN sono i seguenti:

- **Input Layer:** primo layer della rete, ha lo scopo di ricevere i dati grezzi di input (in questo caso le patch con dimensioni  $384 \times 384 \times 3$ ).
- **Convolutional Layer:** è l'elemento costitutivo principale di queste reti ed ha il compito più importante, il processamento dei dati. Il funzionamento si fonda sull'utilizzo di un filtro (o kernel), un array bidimensionale di pesi che scorre sull'immagine ed eseguendo l'operazione di convoluzione. L'output di questo processo è definito mappa delle caratteristiche o mappa di attivazione. I tre iper-parametri fondamentali sono: il numero di filtri, che influisce sulla profondità dell'output, la stride, la distanza di spostamento del kernel sulla matrice di input, e lo zero-padding, utilizzato per un adattamento dimensionale tra filtro e matrice di input.
- **Funzioni di attivazione:** generalmente una Rectified Linear Unity (ReLU), con lo scopo di introdurre la non linearità nel processo di apprendimento; è sempre preceduta da uno strato di normalizzazione che serve a stabilizzare ed accelerare l'allenamento del modello, Batch Normalization (BatchNorm).
- **Pooling Layer:** layer di sotto campionamento che eseguono una riduzione delle dimensioni delle matrici di attivazione, anche questi si basano sull'utilizzo di un filtro che scansiona la matrice, ma rispetto ai layer convoluzionali applica una funzione di aggregazione ai valori

selezionati dal filtro. Le funzioni più comuni sono Max Pooling e Average Pooling, che selezionano il valore più alto, nel primo caso, e la media nel secondo.

- **Fully Connected Layer:** strato in cui avviene la vera e propria classificazione, ogni nodo di output è connesso a un nodo del livello precedente. L'input dello strato sono le matrici di attivazione ridotte a vettori unidimensionali che qui, tramite una funzione di attivazione softmax, vengono classificate con una probabilità che va da 0 a 1.

Il processo di apprendimento di una CNN si basa su alcuni elementi cardine, primi tra tutti i pesi ( $W$ ), che inizializzati casualmente, durante la fase di allenamento devono essere aggiornati affinché l'output sia il più simile possibile a quello reale. L'immagine viene propagata attraverso la rete e i layer descritti fino al layer di output dove viene fornita la predizione (*Forward Propagation*). La predizione e la classificazione reale vengono confrontate e si calcola l'errore attraverso una funzione di perdita (*Loss Function, L*). Lo scopo della rete è la corretta classificazione del input e di conseguenza la minimizzazione della funzione di perdita che quantifica l'errore; per farlo si calcola la derivata della perdita rispetto ai pesi (*Gradiente,  $\frac{\partial L}{\partial W}$* ) e con degli algoritmi di ottimizzazione, che sfruttano la discesa del gradiente, si aggiornano i pesi dei layer in senso opposto al flusso dei dati (*Backward Propagation*) [24]. I pesi vengono, quindi, aggiornati e il processo ripete fino al completamento delle epoche di training previste.

Nel paragrafo 3.3 verranno spiegate le funzioni selezionate e il loro funzionamento più nel dettaglio.

## 3.2 RESNET

Per lo svolgimento del lavoro è stato scelto, come framework di Deep Learning, “Tensorflow”. Questo permette l’utilizzo di numerose CNN, in particolare delle Residual Network o meglio conosciute come ResNet, introdotte da He et al [25], nel 2016.

### 3.2.1 SKIP CONNECTION

Le ResNet introducono un'architettura innovativa basata sull'uso di *skip connections* o collegamenti residui che permettono il flusso diretto delle informazioni tra i livelli della rete bypassandone uno o più, collegando l’output di un layer direttamente a un layer successivo.

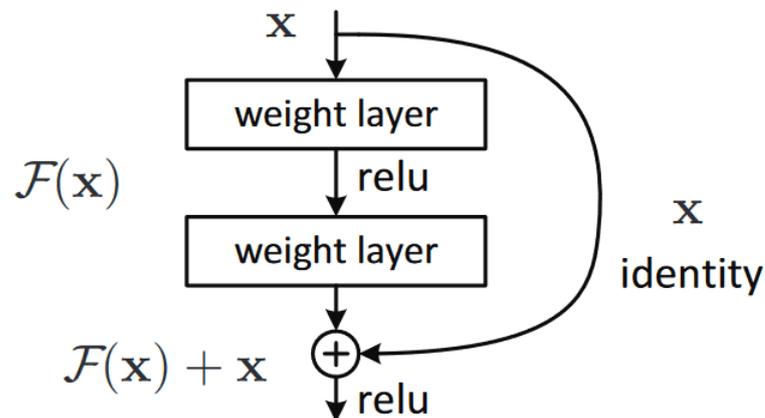


Figure 5: Esempio di Skip connection.

Con  $X$  viene indicato l’input, con  $F(X)$  le trasformazioni non lineari  $F$  effettuate sull’input  $X$  dai layer intermedi. Nelle skip connection l’input è sommato a  $F(X)$  producendo l’output:

$$H(X) = F(X) + X \quad (1)$$

Queste connessioni, che possono essere di tipo diretto o con proiezione, in funzione della dimensione di input e output, servono a migliorare la rete nei seguenti aspetti:

- Evitare il problema della scomparsa del gradiente (*vanishing gradient*, la diminuzione sproporzionata del gradiente durante la retro-propagazione dell’errore, che porterebbe ad un aggiornamento troppo piccolo dei pesi),

- Permette di apprendere nuove informazioni sull'immagine senza perdere però quelle dell'input  $X$ , ricavate dagli strati precedenti [25], e
- Migliorare la convergenza e l'allenamento della rete.

### 3.2.2 ARCHITETTURA RESNET18

La variante scelta è la ResNet18 perché garantisce un buon compromesso tra tempi di addestramento, risultati di classificazione e profondità. Rispetto a modelli più complessi come ResNet-50 o ResNet-101, ResNet-18 presenta un numero inferiore di parametri, riducendo il rischio di overfitting e garantendo tempi di addestramento più contenuti, un aspetto cruciale considerando le limitazioni hardware e conseguenti capacità computazionali disponibili.

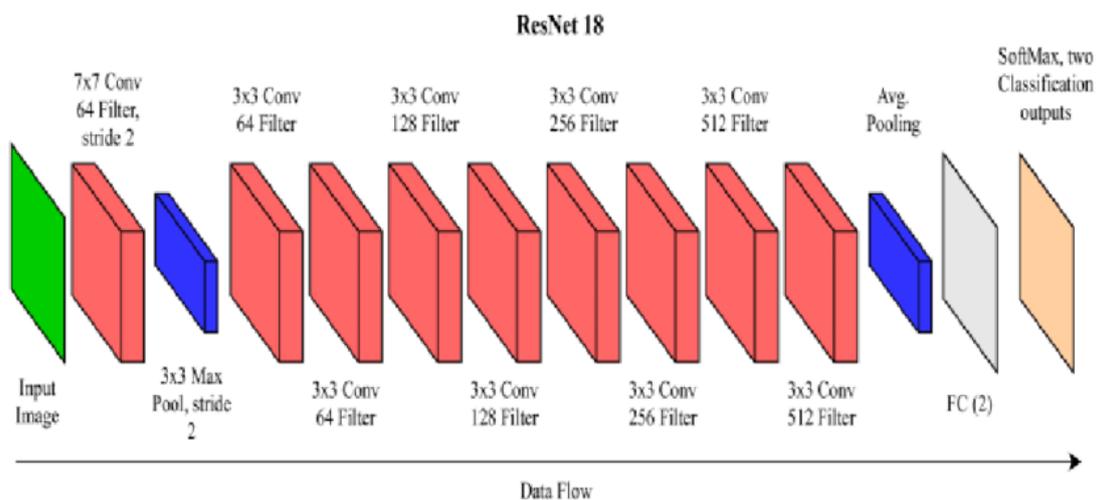


Figure 6: Struttura tipica di una ResNet18.

TensorFlow non presenta nella propria rosa di reti neurali utilizzabili la ResNet18, così è stato utilizzato Keras, l'API ufficiale di alto livello di TensorFlow, che permette di "assemblare" la rete semplicemente richiamando i layer desiderati in modo sequenziale.

Nelle prossime pagine viene riportata l'architettura dettagliata della ResNet18 utilizzata.

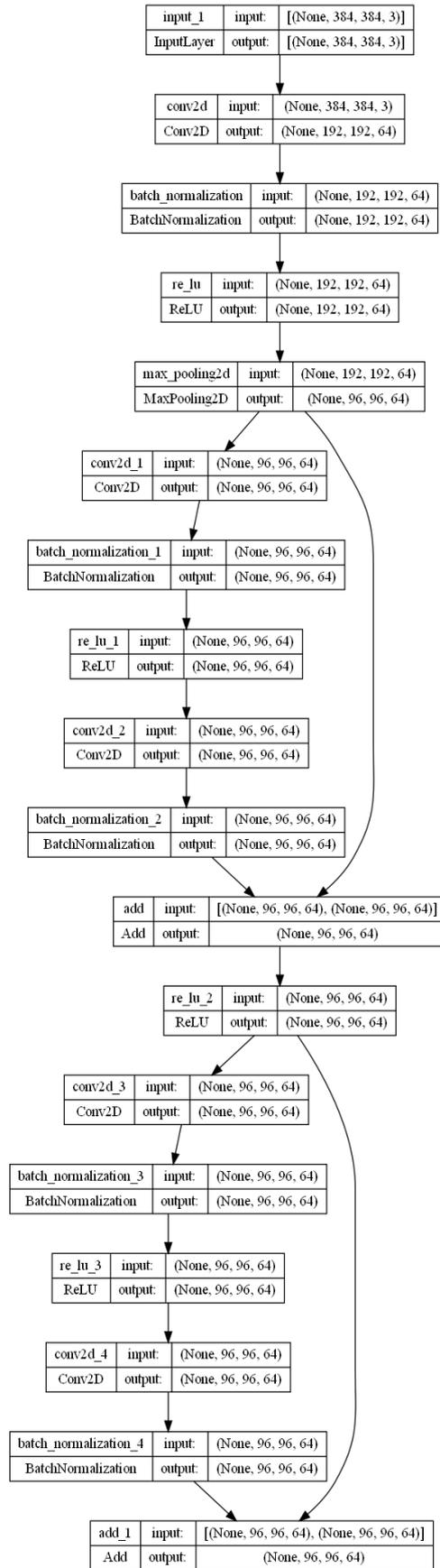


Figure 7: Architettura dettagliata della ResNet18 utilizzata con nomi dei layer e caratteristiche degli stessi. (Parte I)

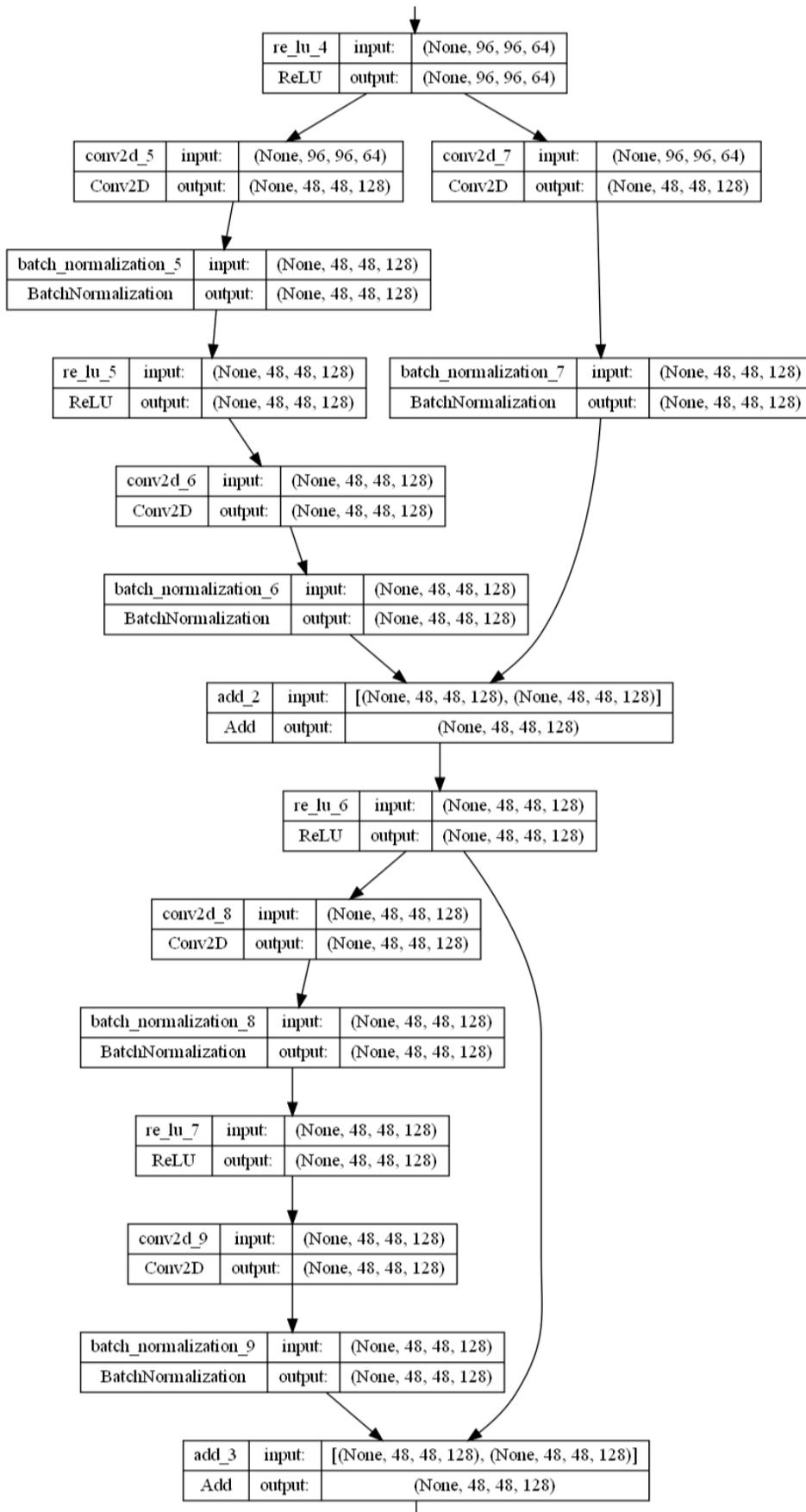


Figure 8: Architettura dettagliata della ResNet18 utilizzata con nomi dei layer e caratteristiche degli stessi. (Parte II)

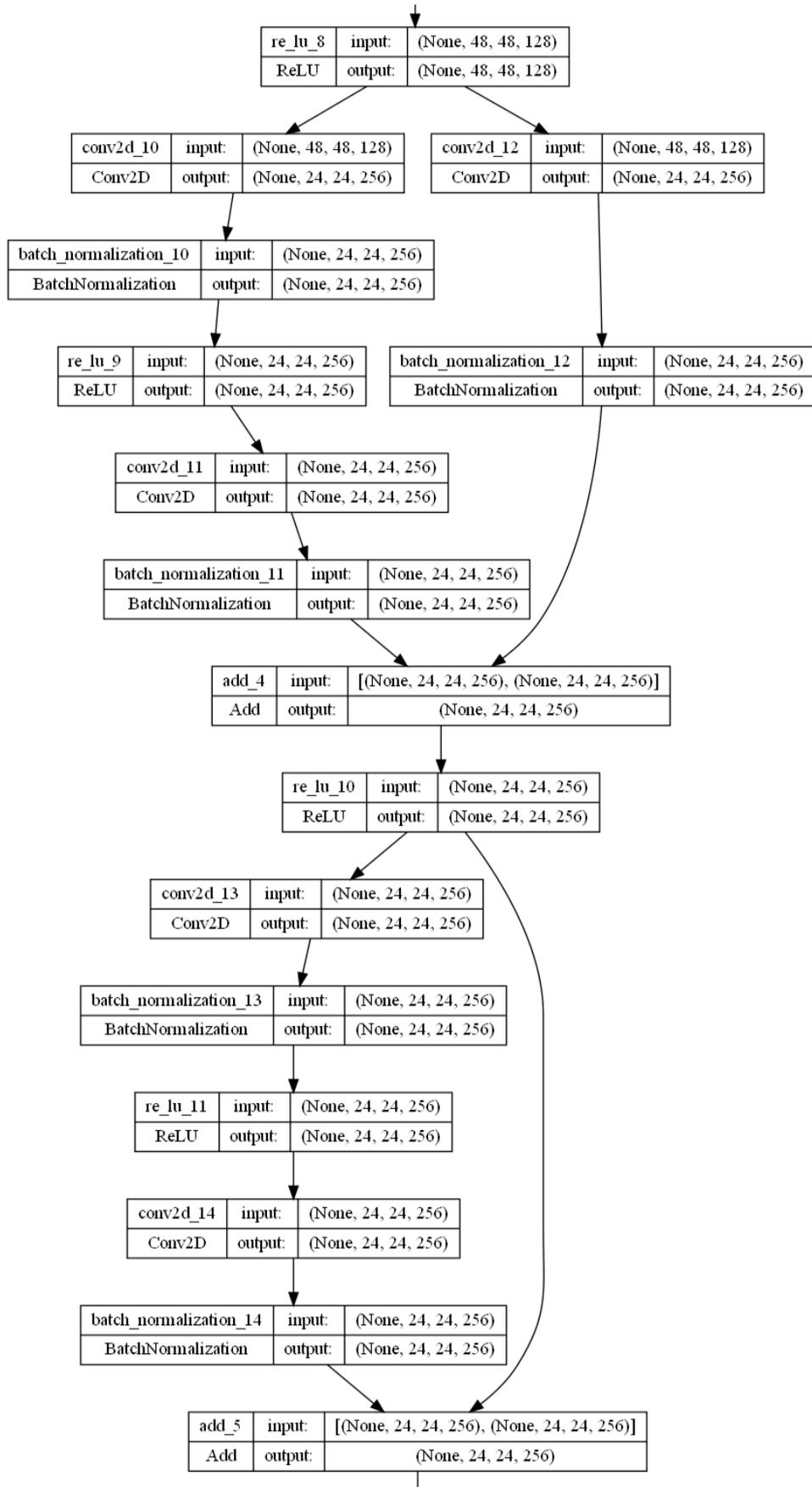


Figure 9: Architettura dettagliata della ResNet18 utilizzata con nomi dei layer e caratteristiche degli stessi. (Parte III)

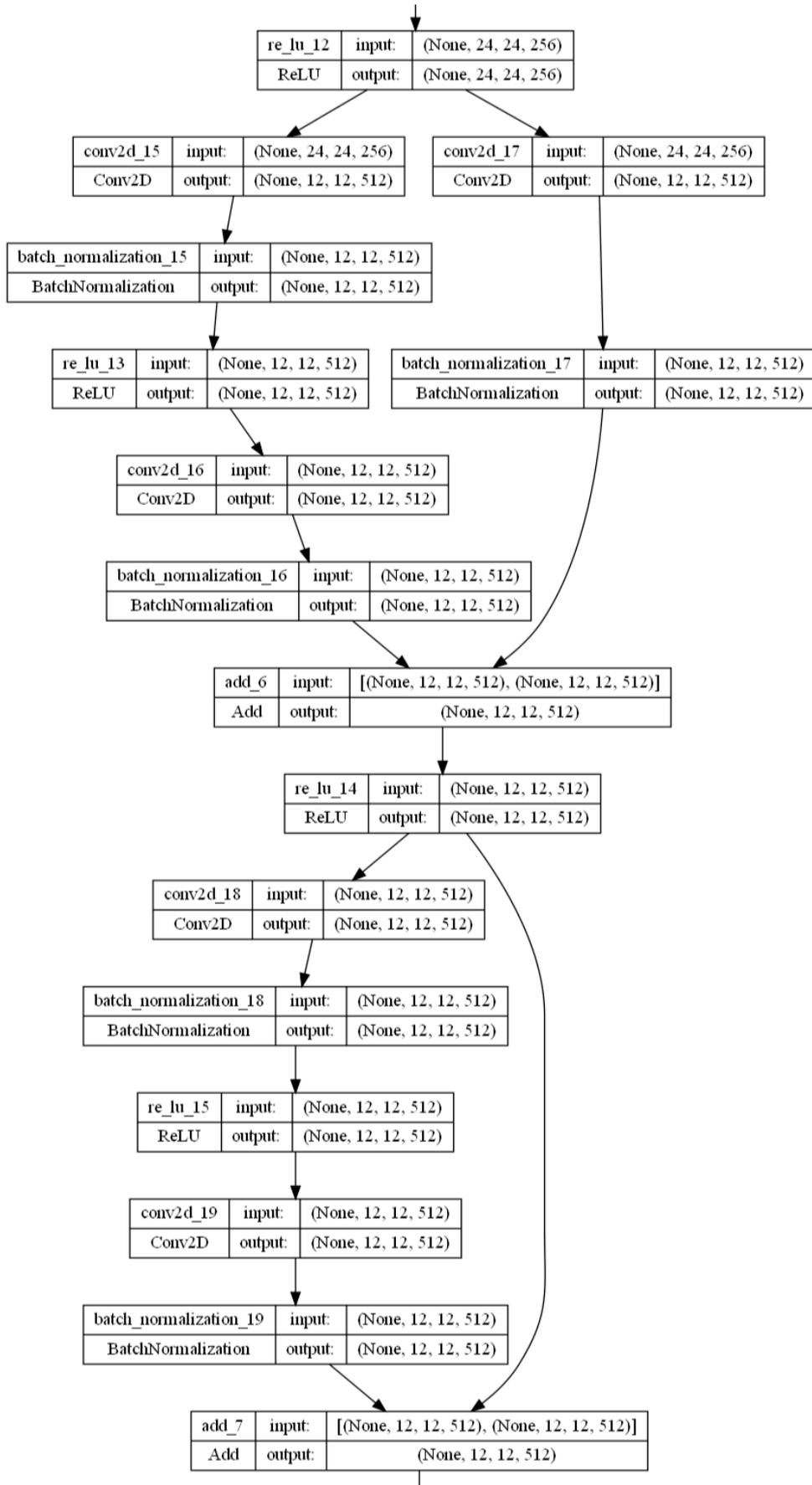


Figure 10: Architettura dettagliata della ResNet18 utilizzata con nomi dei layer e caratteristiche degli stessi. (Parte IV)

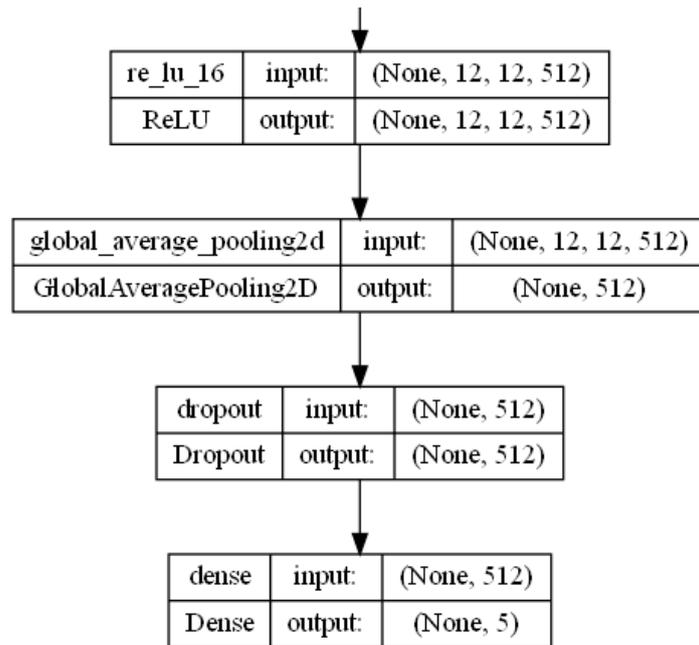


Figure 11: Architettura dettagliata della ResNet18 utilizzata con nomi dei layer e caratteristiche degli stessi. (Parte V)

La ResNet18 generata è caratterizzata da 18 layer allenabili, 17 dei quali convoluzionali (Conv2D), e lo strato completamente connesso (Dense), per un totale di 11,179,077 parametri allenabili.

## 3.3 PARAMETRI DI TRAINING

L'allenamento di una rete è un processo complesso, basato sulla ricerca della combinazione migliore tra le funzioni e i vari parametri scelti per il raggiungimento di buone performance di classificazione; quindi, in questo paragrafo, verranno presentati funzioni e parametri impostati per l'allenamento.

I parametri più immediati e facili da modificare sono:

- **Batch size:** quantità di immagini processate contemporaneamente per un aggiornamento di pesi, impostata pari a 32, buon compromesso tra stabilità e memoria disponibile.
- **Numero di epoche:** Quantità di volte che il training set, diviso in batch, attraversa la rete.

### 3.3.1 DROPOUT

Le reti neurali durante l'allenamento imparano anche il rumore statistico, in particolare ogni neurone impara in modo tale da correggere i possibili errori commessi da altre unità, ciò genera complessi co-adattamenti; di conseguenza la capacità di generalizzare su immagini del dataset non viste viene compromessa. Questo problema è definito overfitting [26]. Per limitare l'effetto dell'overfitting esistono diverse tecniche di regolarizzazione (*Paragrafo 3.3.2*) che penalizzano i pesi della rete e il Dropout; come suggerisce il nome, questa tecnica permette durante la fase di training della rete di 'spegnere momentaneamente' alcuni neuroni. La rimozione di un neurone porta alla rimozione di tutte le sue connessioni in entrata, in uscita e i pesi che applica come mostrato in *Figura 12* [27], ad ogni iterazione la configurazione dei neuroni selezionati per essere spenti cambia e di conseguenza la struttura delle connessioni.

La probabilità selezionata nel presente lavoro che ciascun neurone venga spento è  $p = 40\%$  e come si vede in *Figura 11* è stato posizionato prima dello strato completamente connesso.

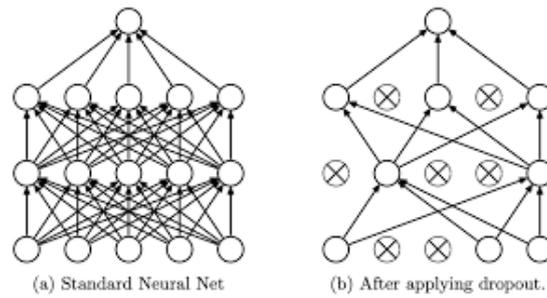


Figure 12: Esempio grafico del dropout e dell'eliminazione dei neuroni e delle loro connessioni durante la fase di training.

### 3.3.2 REGOLARIZZAZIONE $L_2$

Nello strato completamente connesso, oltre alla tecnica del dropout, è stata utilizzata la funzione di regolarizzazione  $L_2$  conosciuta anche come 'weight decay', che aggiunge una penalizzazione al valore dei pesi della rete per evitare che diventino troppo grandi; permette di:

- ridurre la complessità del modello
- evitare l'overfitting e
- stabilizzare il training.

$$L = L + \lambda \cdot \sum W^2 \quad (4)$$

Dove:

- $L$  è la funzione di perdita,
- $\lambda$  è il fattore di regolarizzazione
- $W$  è il valore dei pesi della rete.

### 3.3.3 OTTIMIZZATORE - SGD

L'ottimizzatore scelto per l'aggiornamento dei pesi della rete è il *Stochastic Gradient Descent* (SGD), uno degli algoritmi di ottimizzazione più utilizzati nell'addestramento delle reti neurali.

$$W = W - \eta \cdot \frac{\partial L}{\partial W} \quad (5)$$

Dove:

- $W$  è il peso della rete neurale
- $\eta$  è il *learning rate*, che controlla la grandezza dell'aggiornamento
- $\frac{\partial L}{\partial W}$  è il gradiente della funzione di perdita rispetto a  $W$

Nell'implementazione di tale algoritmo si è scelto di applicare la variante con *Nestorov Momentum* che esegue un controllo anticipato sul gradiente, rendendolo più efficiente, meno rumoroso e più stabile perché in grado di superare i minimi locali.

$$W = W - \left( \gamma v + \eta \cdot \frac{\partial L}{\partial W'} \right) \quad (5.1)$$

Dove:

- $v$  è una memoria della velocità di aggiornamento, per evitare le oscillazioni
- $\gamma$  controlla quanto il valore passato influisca l'aggiornamento attuale

Il *learning rate* scelto è  $\eta = 10^{-4}$  e il *momentum*  $\gamma = 0.9$ .

### 3.3.4 LOSS FUNCTION - CROSS ENTROPY LOSS

Per misurare la deviazione tra la classe vera e quella predetta, la funzione di perdita scelta è la *Cross Entropy Loss*, una funzione di perdita molto comune nei task di classificazione di immagini multi-classe con etichette in formato one-hot (es. Classe 0  $\rightarrow$  [1, 0, 0, 0, 0], Classe 3  $\rightarrow$  [0, 0, 0, 1, 0]).

La *Cross-Entropy* misura la distanza tra due distribuzioni di probabilità:

- La distribuzione delle etichette vere (one-hot)
- La distribuzione delle predizioni del modello (output softmax)

$$L = - \sum_{i=1}^C y_i \log(\hat{y}_i) \quad (2)$$

Dove:

- $C$  è il numero di classi (5 in questo task)
- $y_i$  è il valore target (1 per la classe corretta, 0 per le altre)
- $\hat{y}_i$  è la probabilità predetta dal modello per la classe  $i$  (output softmax)

In particolare, a causa dello sbilanciamento del dataset, con classi più rappresentate di altre, per dare più valore a quelle meno rappresentate, per questo lavoro di tesi, è stata utilizzata una variante della *Cross-Entropy Loss* che introduce il peso delle classi, *Weighted Cross-Entropy Loss*.

I pesi per ciascuna classe sono stati calcolati utilizzando il metodo dell'inverso della frequenza:

$$w_i = \frac{N}{n_i} \quad (3)$$

Dove:

- $w_i$  è il peso della classe
- $N$  è il numero totale di immagini del Training set
- $n_i$  è il numero di occorrenze dei campioni della classe  $i$

L'Equazione 2 con l'aggiunta del parametro  $w_i$  diventa:

$$L = - \sum_{i=1}^c w_i \cdot y_i \log(\hat{y}_i) \quad (2.1)$$

I pesi  $w_i$  applicati all'Equazione 2.1 sono i seguenti:

<i>Classe</i>	<i>Peso</i>
<i>0</i>	<i>18.15</i>
<i>1</i>	<i>34.81</i>
<i>2</i>	<i>3.50</i>
<i>3</i>	<i>3.20</i>
<i>4</i>	<i>3.13</i>

Tabella 6: Pesi delle classi applicate alla Weighted Cross-Entropy Loss Function.

### 3.4 CRITERI DI VALUTAZIONE

Per valutare l'andamento della fase di training e dei risultati alla fine dello stesso, sono stati usati diversi criteri di valutazione; lo strumento più utile per una visualizzazione grafica e qualitativa è la *Confusion Matrix*, che permette la valutazione delle prestazioni, confrontando le classi reali e le classificazioni operate dalla rete sul Validation set, ottenendo una quantificazione degli errori commessi e la loro distribuzione tra le classi.

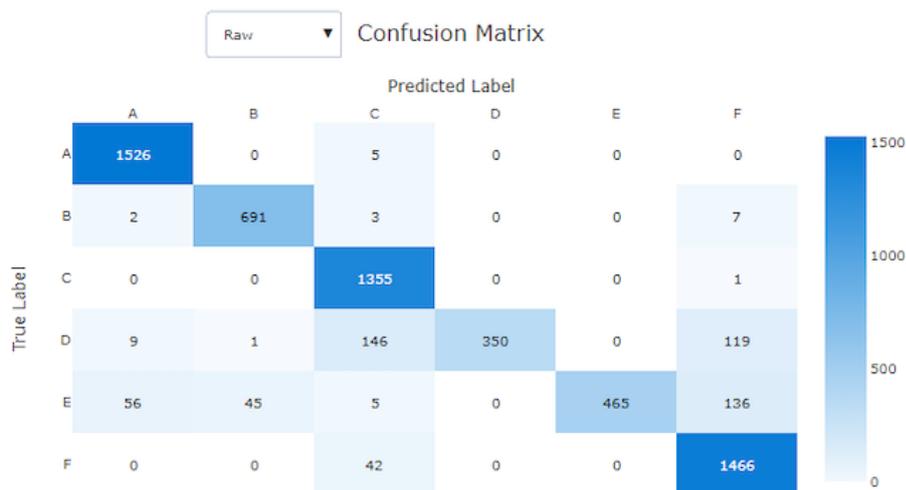


Figure 13: Esempio di Confusion Matrix.

Dalle Confusion Matrix, grazie al pattern di classificazione che si viene a creare, è possibile estrarre delle metriche numeriche:

$$Accuracy = \frac{\sum TP}{\sum \text{tutti i valori della matrice}} \quad (6)$$

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \quad (7)$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \quad (8)$$

Dove:

- TP, sono i veri positivi, immagini correttamente classificate
- FP, sono i falsi positivi, immagini di altre classi classificate come classe  $i$
- FN, sono i falsi negativi, immagini di classe  $i$  classificate come altre classi

Le *Equazioni 7 e 8* descrivono meglio la qualità della classificazione rispetto alla  $\delta$  che soffre di un bias dovuto dallo sbilanciamento delle classi.

Inoltre, durante la fase di training sono stati considerati fondamentali i valori della funzione di perdita calcolate su Training Set e Validation Set (*val\_loss*).

# 4 HEATMAP

---

Lo scopo di questo lavoro di tesi è fornire uno strumento di supporto alla classificazione dei tessuti del surrene e in particolare la diagnosi tra ACA e ACC; il modo migliore è rendere facilmente comprensibile e interpretabile l'output fornito agli operatori sanitari. Per farlo si è deciso di creare delle mappe calore o heatmap per evidenziare la classificazione delle Tiles.

## 4.1 CREAZIONE HEATMAP

Il processo seguito per la classificazione di una Tile alla volta può essere descritto con i seguenti punti:

- I. Suddivisione in patch:** Il classificatore, come spiegato in precedenza, lavora e processa immagini con formato RGB di dimensioni  $384 \times 384 \times 3$ ; quindi, proprio come fatto per la creazione del dataset di allenamento (*Paragrafo 2.1*) è necessario dividere le immagini originali ( $1000 \times 1000 \times 3$ ) in patch. La dimensione della Tile non è divisibile per la grandezza delle patch, quindi, è stata utilizzata la tecnica del padding (aggiunta di una quantità specifica di pixel, in questo caso 76 per lato) riflettendo il bordo originale, in modo tale da poter classificare nella sua interezza la Tile.

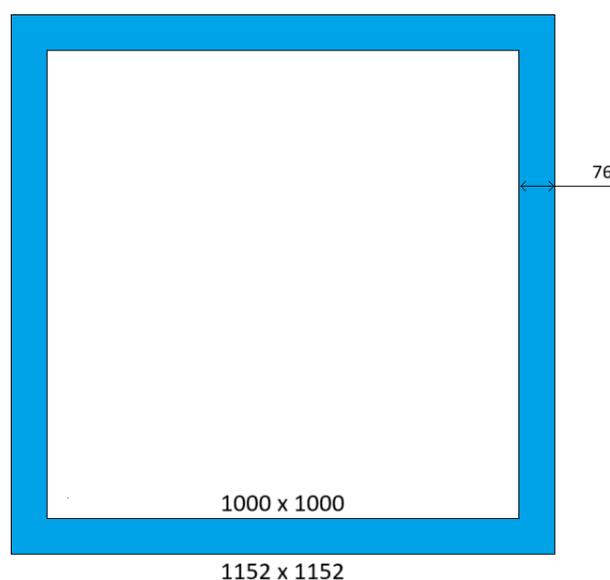
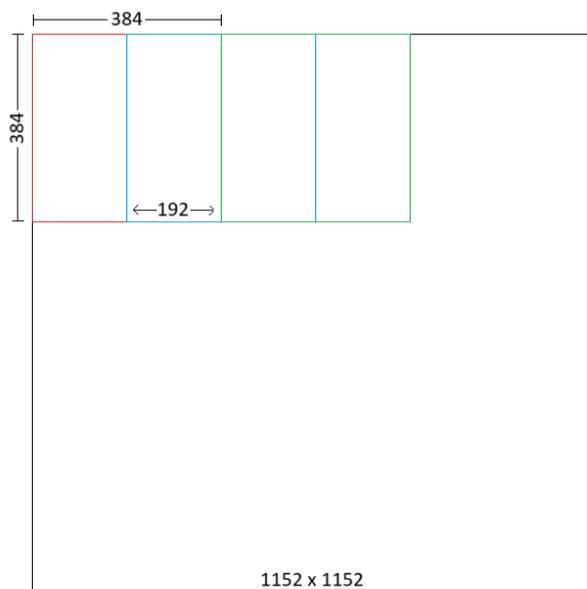


Figure 14: Visualizzazione della tecnica del Padding.

La dimensione delle “nuove immagini” è (1152x1152x3). Le immagini sono state estratte con una sovrapposizione, in entrambe le direzioni, verticale e orizzontale, pari al 50%, in modo tale che la ResNet18 allenata classificasse più volte la stessa porzione della Tile, con lo scopo di rendere più solida la classificazione finale. Con il metodo descritto si ottengono 25 patch per ciascuna Tile.



*Figure 15: Visualizzazione della Patch Extraction.*

- II. Classificazione:** Le patch, estratte nel passo precedente, vengono normalizzate e fornite in input alla ResNet18 pre-allenata che le classifica. Si ottiene, così, un vettore con le classi predette con dimensione pari al numero di patch in input.
- III. Majority voting:** Utilizzando le coordinate di ciascuna patch, si individuano le porzioni della Tile formate da un'intersezione di quattro patch (dimensioni 192x192), si conta la frequenza delle classi predette per le patch prese in considerazione e la classe più frequente tra quelle predette viene assegnata alla zona di intersezione.

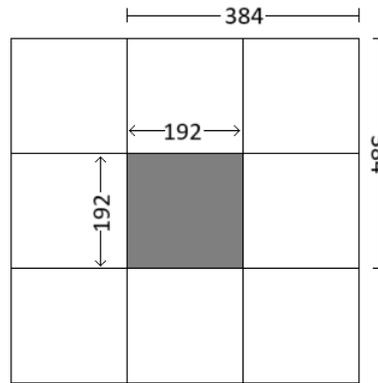


Figure 16: Visualizzazione della sovrapposizione delle patch e della zona di intersezione soggetta a classificazione per majority voting.

- IV. Maschera in scala di grigi:** Inizializzata come immagine nera (valore di grigio = 0), viene riempita utilizzando il livello di grigio della classe assegnata alle aree di intersezione tra quattro patch. Il risultato sarà una maschera in scala di grigi, con un artefatto di bordo di 192 pixel neri (non classificati), si elimina il padding iniziale di 76 pixel, riducendo l'artefatto a 116 pixel non classificati. Per ovviare a questo problema e fornire una heatmap che classifichi interamente la Tile è stata utilizzata la stessa tecnica di padding utilizzata precedentemente, in modo tale da assegnare una colorazione (e quindi una classificazione) per riflessione pari a quella interna.
- V. Heatmap:** La maschera in scala di grigi è stata trasformata, tramite l'applicazione di un filtro Gaussiano, per lo *smoothing* dei bordi tra classi, e di una colormap, (cv2.COLORMAP\_JET) in una Heatmap.

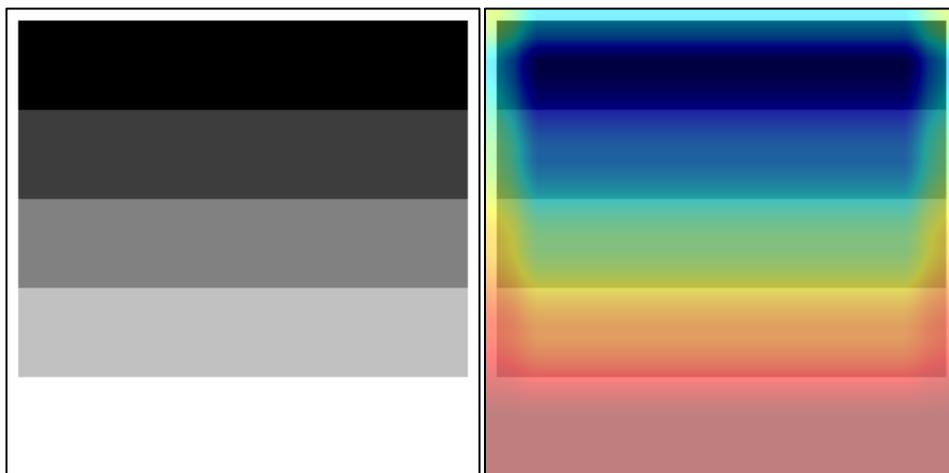


Figure 17: Confronto tra la Maschera in scala di grigi (a) e la Heatmap (b) ottenuta con l'applicazione della colormap 'JET'.

La Heatmap, infine, è stata sovrapposta alla Tile integrale, sfruttando il canale di trasparenza  $\alpha$  delle immagini RGBA, fornendo una visualizzazione complessiva più comoda e semplice.

## 4.2 CRITERI DI VALUTAZIONE DELLE HEATMAP

Per poter effettuare un confronto bisogna determinare dei termini di paragone, sono state usate le heatmap nella loro forma grezza, ottenute seguendo i passi dal I al IV con le Tile contenute nel Test set, sulle quali la rete non è stata allenata.

Le metriche comunemente usate per valutare l'errore tra due immagini sono le seguenti:

- **Mean Absolute Error Pixel-wise:** Calcola l'errore medio assoluto tra i pixel di due immagini, è una misura della differenza tra le due.

$$MAE = \frac{1}{N} \sum_{p=1}^N |l_p - h_p| \quad (9)$$

Dove:

- $N$  è il numero di pixel totale nell'immagine,
  - $l_p$  è il valore del pixel  $p$  nell'immagine originale (*label*)
  - $h_p$  è il valore del pixel  $p$  nella maschera in scale di grigi (*heatmap grezza*).
- **Root Mean Square Error Pixel-wise:** è la radice quadrata dell'errore quadratico medio tra pixel, penalizza maggiormente gli errori grandi rispetto all'*Equazione 9*.

$$RMSE = \frac{1}{N} \sqrt{\sum_{p=1}^N (l_p - h_p)^2} \quad (10)$$

L'uso combinato delle due metriche fornisce una visione più completa dell'errore commesso e permette di stimare quanto ciò che è stato predetto dal classificatore sia simile alla label manuale.<sup>4</sup>

---

<sup>4</sup> Si noti che le metriche proposte sono applicabili a un confronto tra due immagini, di conseguenza nel *Paragrafo 5.2* verrà presentata una media dei valori ottenuti per tutte le immagini del dataset.

# 5 RISULTATI

---

## 5.1 RISULTATI ALLENAMENTO

La fase di training del modello è stata condotta utilizzando un controllo sul valore della funzione di perdita calcolata sul Validation set, in modo tale che ogni qualvolta si registrasse un minimo, quindi una quantità di errori bassa, fosse salvata la configurazione del modello, compresi i pesi adottati per quell'epoca.

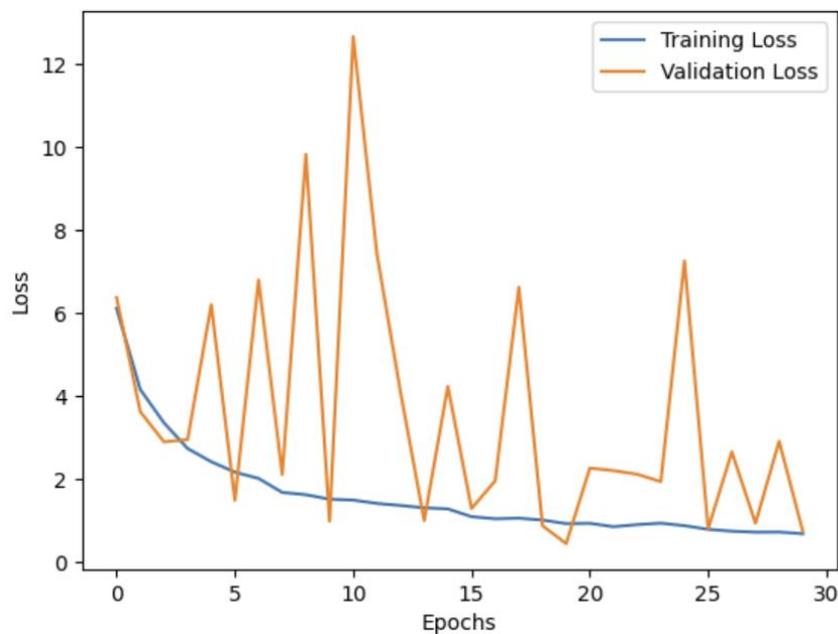


Figure 18: Valori della funzione di perdita sul validation set.

Il valore minimo di 'val\_loss' registrato è stato pari a 0.45, per cui questa è stata la configurazione della ResNet18 scelta per la fase di testing, calcolo delle prestazioni e, soprattutto, per la creazione delle heatmap secondo il metodo proposto.

		Predette					
		0	1	2	3	4	
Reali	0	3	0	3	4	2	12
	1	1	5	0	11	0	17
	2	5	4	71	8	30	118
	3	27	0	0	802	46	875
	4	90	6	9	9	745	859
							1881

Tabella 7: Confusion Matrix sul test set.

L'Accuracy calcolata con l'Equazione 6 è:

$$Accuracy = 86,4 \%$$

I valori di Precision e Recall per classe calcolati con le Equazioni 7 e 8 sono riassunti nella tabella seguente:

Classe	Precision <sub>i</sub>	Recall <sub>i</sub>
0	2,3 %	25 %
1	33,3 %	29,4 %
2	85,5 %	60,2 %
3	96,2 %	91,7 %
4	90,5 %	86,7 %

Tabella 8: Valori di precision e recall per classe.

## 5.2 RISULTATI HEATMAP

Come anticipato nel *Paragrafo 4.2*, la bontà di classificazione e rappresentazione delle predizioni sulle Tiles è stata calcolata utilizzando le maschere in scala di grigi ottenute al passo IV. Nelle *Figure* seguenti viene riportato un esempio per classe.



*Figure 19: Risultato parziale di classificazione di una Tile NAM; da sinistra a destra: Tile originale, label manuale e Maschera in scala di grigi.*



*Figure 20: Risultato parziale di classificazione di una Tile NAC; da sinistra a destra: Tile originale, label manuale e Maschera in scala di grigi.*

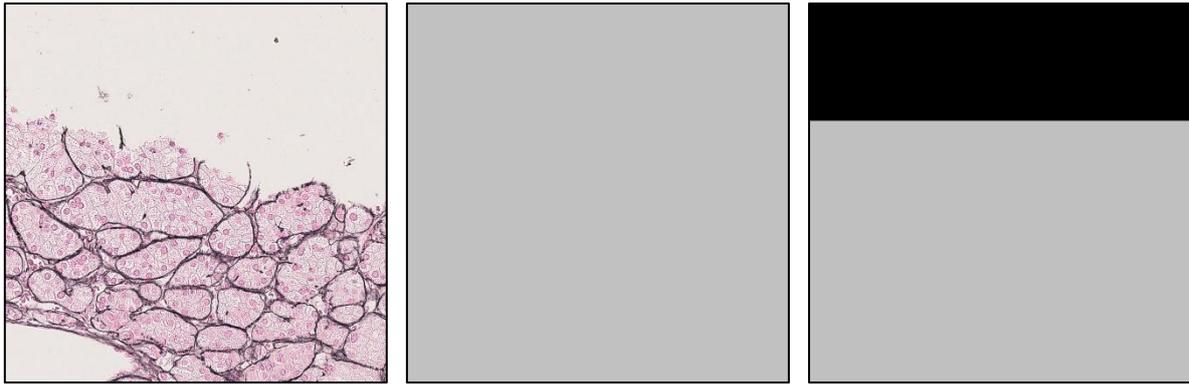


Figure 21: Risultato parziale di classificazione di una Tile ACC, da sinistra a destra: Tile originale, label manuale e Maschera in scala di grigi.

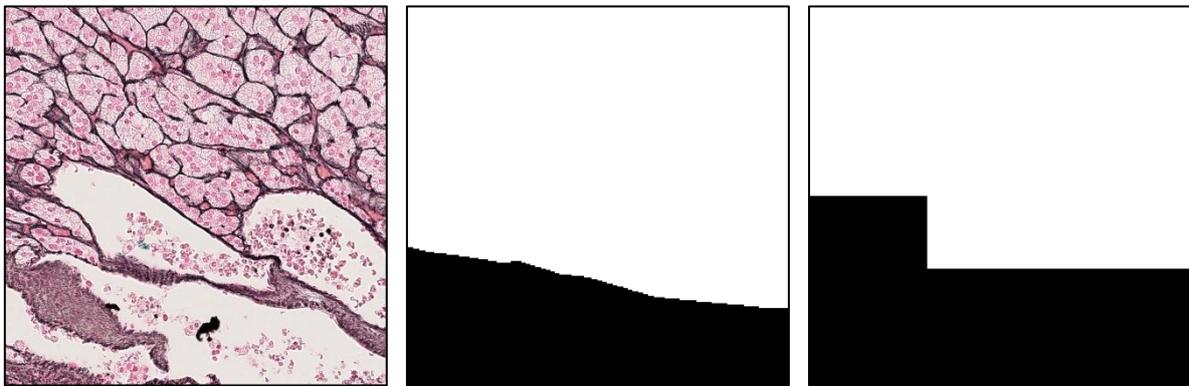


Figure 22: Risultato parziale di classificazione di una Tile ACA; da sinistra a destra: Tile originale, label manuale e Maschera in scala di grigi.

I valori medi calcolati sull'intero Test Set di MAE e RMSE con le *Equazioni 9 e 10* sono i seguenti:

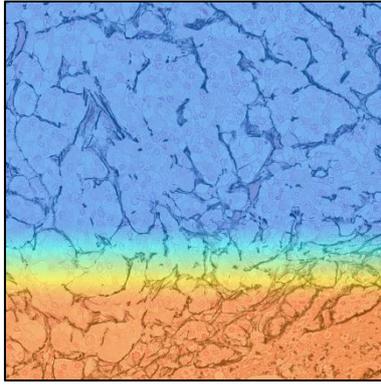
$$mean\_MAE = 0,103$$

$$mean\_RMSE = 0,158$$

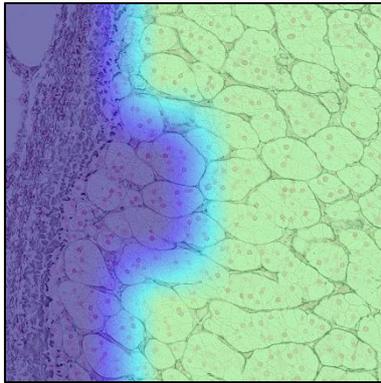
L'output finale del metodo presentato nel lavoro di tesi è presentato dalle seguenti heatmap, ottenute dalla sovrapposizione delle maschere alle Tile originali.<sup>5</sup>

---

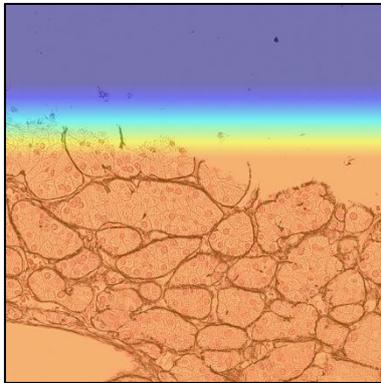
<sup>5</sup> Per continuità dimostrativa le heatmap proposte sono quelle ricavate dalle maschere presentate nelle figure da 19 a 22.



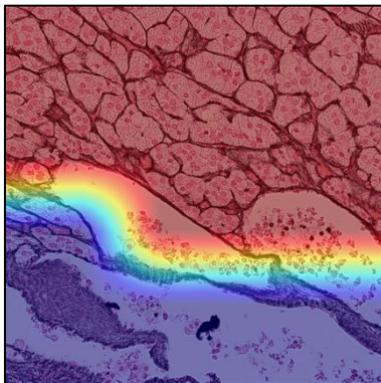
*Figure 23: Heatmap della Tile NAM.*



*Figure 24: Heatmap della Tile NAC.*



*Figure 25: Heatmap della Tile ACC.*



*Figure 26: Heatmap della Tile ACA.*

## 6 CONCLUSIONI

---

Il metodo proposto ha ottenuto complessivamente buone performance di classificazione, con un'accuracy media dell'86%. Tuttavia, un'analisi più dettagliata dei risultati (*Tabella 8*) evidenzia una forte variabilità delle percentuali di classificazione a seconda della classe. Questo sbilanciamento riflette la distribuzione iniziale del dataset, influenzando le capacità della ResNet-18 di predire correttamente alcune classi.

Nonostante le modifiche apportate per riequilibrare il dataset (*Tabella 5*), in particolare nel training set, dove le classi NAC, ACA e ACC presentano ora una distribuzione più uniforme, le classi NAM e OT rimangono numericamente sottorappresentate a causa della loro scarsa presenza nel dataset originale. Le strategie adottate durante l'addestramento, tra cui dropout, regolarizzazione e l'uso di pesi per le classi, hanno migliorato l'apprendimento della rete, ma non hanno portato alla capacità di generalizzazione desiderata e il problema dell'overfitting non è stato risolto, come si nota anche dalle oscillazioni del valore di `val_loss` (*Figura 18*).

Le performance di classificazione sulle Tile rispecchiano i risultati ottenuti sul test set, come evidenziato dagli errori calcolati. Tuttavia, l'output finale, pensato per la visualizzazione da parte dei medici sotto forma di mappe di calore, si è rivelato uno strumento utile per una lettura più intuitiva e un'interpretazione rapida dei risultati.

In conclusione, il metodo presentato in questo lavoro di tesi rappresenta una solida base di partenza per futuri sviluppi. Un miglior bilanciamento delle classi nel dataset, unito a un'ottimizzazione più mirata dei parametri, potrebbe migliorare ulteriormente la capacità di generalizzazione della rete e portare a performance di classificazione ancora più significative.

## 7 RIFERIMENTI

---

- [1] W. C. K. S. e. a. Megha R, *Anatomy, Abdomen and Pelvis: Adrenal Glands (Suprarenal Glands)*, Treasure Island (FL): StatPearls Publishing, [Updated 2022 Oct 17].
- [2] K. A. Y. J. A. W. E. J. B. P. D. H. K. O. K. J. E. J. M. W. P. D. J. Gordon Betts, «Anatomy and Physiology 2e,» *OpenStax*, Apr 20, 2022.
- [3] L. Crumbie, «Adrenal glands anatomy,» Last reviewed: October 30, 2023. [Online]. Available: <https://www.kenhub.com/en/library/anatomy/adrenal-glands>.
- [4] W. C. J. I. Dutt M, *Physiology, Adrenal Gland.*, vol. StatPearls [Internet], Treasure Island (FL): StatPearls Publishing, [Updated 2023 May 1].
- [5] L. R. Martin, «The Adrenal Glands,» Last updated April 8, 2024. [Online]. Available: <https://teachmeanatomy.info/abdomen/viscera/adrenal-glands/>.
- [6] J. L. E. B. M. R. Y. M. GEORG MANSMANN, «The Clinically Inapparent Adrenal Mass: Update in Diagnosis and Management,» *Endocrine Reviews*, 2004.
- [7] T. K. A. C. S. A. R. V. M. K. A. C. E. M. J. S. M. B. S. G. T. J. & H. G. D. Else, «Adrenocortical carcinoma,» *Endocrine reviews*, 2014.
- [8] B. A. a. M. Fassnacht, «CLINICAL REVIEW: Adrenocortical Carcinoma: Clinical update,» *The Journal of Clinical Endocrinology & Metabolism*, 2006.
- [9] R. S. B. F. E. M. E. C. A. O.-F. E. Michalkiewicz, «Clinical and Outcome Characteristics of Children With Adrenocortical Tumors: A Report From the International Pediatric Adrenocortical Tumor Registry,» *Journal of clinical oncology*, 2004.
- [10] M. L. R. D. E. V. M. B. J. & T. F. Papotti, «The Weiss score and beyond--histopathology for adrenocortical carcinoma,» *Hormones & cancer*, 2(6), 333–340, 2011.
- [11] M. S. A. A. A. F. S. L. P. D. R. & S. P. M. Sherlock, «Adrenal Incidentaloma,» *Endocrine reviews*, 41(6), 775–820., 2020.

- [12] S. C. V. & B. A. Grisanti, «Terapia del carcinoma corticosurrenalico dopo fallimento del mitotane.,» *L'Endocrinologo* 24, 533–537., 2023.
- [13] D. E, F. A, V. M, N. G, S. R, G. G, C. R, D. C. P, V. L, G. M, D. T. AP, L. P, M. M, M. F, B. A, T. M e P. M., «The Reticulin Algorithm for Adrenocortical Tumor Diagnosis: A Multicentric Validation Study on 245 Unpublished Cases.,» *The American Journal of Surgical Pathology*, 2013.
- [14] C. S. K. J. K. S. J. C. C. J. K. S. K. Y. Kim TM, «Fully automatic volume measurement of the adrenal gland on CT using deep learning to classify adrenal hyperplasia.,» *European Radiology*., vol. 33, n. 6, pp. 4292-4302, 2023 Jun.
- [15] N. T. N. A. Y. Y. K. K. E. M. O. Y. I. K. Kusunoki M, «A deep learning-based approach for the diagnosis of adrenal adenoma: a new trial using CT.,» *The British journal of radiology*., 2022 Jul 1.
- [16] Z. S. K. S. F. D. E. T. Y. I. B. & B. J. E. Yashbir Singh, «Deep learning approach for differentiating indeterminate adrenal masses using CT imaging.,» *Abdom Radiol*, vol. 48, p. 3189–3194, 2023.
- [17] P. F. C. H. Y. D. J. X. C. W. J. M. Y. C. Y. Y. L. L. C. X. D. X. G. & S. F. Alimu, «Artificial intelligence with a deep learning network for the quantification and distinction of functional adrenal tumors based on contrast-enhanced CT images.,» *Quantitative imaging in medicine and surgery*, vol. 13(4), p. 2675–2687, 2023.
- [18] X. G. X. C. C. Z. Y. Z. Z. L. M. L. P. Y. A. L. X. L. L. C. B. T. & Z. C. Yi, «Adrenal incidentaloma: machine learning-based quantitative texture analysis of unenhanced CT can effectively differentiate sPHEO from lipid-poor adrenal adenoma.,» *Journal of Cancer*, vol. 9(19), p. 3577–3582, 2018.
- [19] J. N. a. K. J. a. C. P. a. e. a. Kather, «Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study.,» *PLOS Medicine*, vol. 16, n. 1, p. e1002730, 2019.
- [20] N. O. P. S. S. T. N. N. S. M. F. D. M. A. L. R. N. & T. A. Coudray, «Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning.,» *Nature Medicine*, vol. 24, n. 10, pp. 1559--1567, 2018.

- [21] M. A. J. Z. W. L. L. L. Xuebin Xu, «A High-Precision Classification Method of Mammary Cancer Based on Improved DenseNet Driven by an Attention Mechanism,» *Comput Math Methods Med*, 2022.
- [22] J. W. R. H. H. e. a. Dolezal, «Deep learning generates synthetic cancer histology for explainability and education.,» *npj Precision Oncology*, Vol. %1 di %27, 49, 2023.
- [23] K. M. Y. M. Y. K. Masakazu Matsugu, «Subject independent facial expression recognition with robust face detection using a convolutional neural network.,» *Neural Networks 16*, p. 555–559, 2003.
- [24] A. Provino, «Neural Network Training Theory,» 3 marzo 2020. [Online]. Available: <https://www.andreaprovino.it/neural-network-training-6s191>.
- [25] X. Z. S. R. a. J. S. K. He, «Deep Residual Learning for Image Recognition,» *2016 IEEE Conference on Computer Vision and Pattern Recognition*.
- [26] H. Yadav, «Dropout in Neural Networks,» 5 Jul 2022. [Online]. Available: <https://towardsdatascience.com/dropout-in-neural-networks-47a162d621d9/>.
- [27] G. H. A. K. I. S. R. S. Nitish Srivastava, «Dropout: A Simple Way to Prevent Neural Networks from Overfitting,» *Journal of Machine Learning Research 15*, pp. 1929-1958, 2014.



# RINGRAZIAMENTI

---

Finalmente, dopo anni, si è concluso questo capitolo della mia vita e sono pronto ad aprirne uno nuovo. Chi mi conosce sa che non sono mai stato bravo né a parlare né ad esprimere le mie emozioni, ma si dà il caso che scrivere dei ringraziamenti sia una consuetudine alla fine del percorso universitario... quindi, eccomi qui a provarci.

Un primo ringraziamento speciale va ai miei genitori, alla costante e rumorosa presenza di mamma e a quella più silenziosa ma non meno amorevole di papà, senza di loro probabilmente non sarei qui in questo momento, sia come uomo che come ingegnere. Spero di aver ricambiato gli anni di sacrifici con questo preciso momento di felicità e orgoglio.

Ringrazio tutti i miei familiari, (nonni, zii, cugini ecc.) dai quali mi sono dovuto allontanare per raggiungere questo obiettivo e che mi hanno sempre supportato in tutte le mie scelte.

Ringrazio gli amici di sempre, il gruppo “sobrietà”: Agnese, Andrea, Antonio, Betta e Silvia, divisi dai percorsi di vita che abbiamo scelto, ma mai lontani, con loro ho condiviso i momenti migliori e peggiori di questo percorso, attraverso uno schermo, programmando la prossima rimpatriata o l’ennesima estate da vivere insieme. Tra questi un ringraziamento speciale va a Musu (non me ne vogliano gli altri), che, con il suo modo di essere sempre presente con il sorriso in bocca, con la battuta giusta al momento giusto, è diventato un fratello.

Ringrazio Alessandra, Luigi, Maddalena, Suchet e Valeria, coloro che ho definito casa e famiglia quando per la prima volta in vita mia ero lontano da queste, con loro ho mosso i primi passi in una nuova città e nella vita universitaria; i pranzi al Cavour, gli aperitivi alle Panche, i pomeriggi a studiare in aula studio rimarranno un elemento indelebile di questo percorso.

Ringrazio anche tutti coloro che in questo periodo hanno fatto parte della mia vita in un modo o nell’altro, amici e colleghi, su tutti Anna e Michele. Menzione particolare in questo caso spetta a Giorgia, “ ‘a magghioia”, che non meriterebbe nemmeno di stare in questa categoria per la valenza che ha assunto in così breve tempo. Tra lezioni, studio e palestra è diventata un punto di riferimento, in poche parole la persona più bella che questo percorso potesse lasciarmi.

Infine, non per importanza, ringrazio Riccardo che in questi anni, da fratello si è trasformato in coinquilino e amico. L’amore fraterno è complicato da spiegare e soprattutto da mettere nero su

bianco, per questo a differenza degli altri a lui non dedico un paio di righe in questo foglio, ma la promessa che per lui sarò sempre il fratello maggiore che merita.

A voi tutti, presenti e non...

Grazie, vi voglio Bene.