

UNIVERSITÀ DEGLI STUDI DEL MOLISE
DIPARTIMENTO DI MEDICINA E SCIENZE DELLA SALUTE
“VINCENZO TIBERIO”



Corso di Dottorato in Medicina Traslazionale e Clinica

Tesi Sperimentale in Radiomica

RADIOMICA, METODI FORMALI E TC AD ALTA RISOLUZIONE
PER LA DIAGNOSI AUTOMATIZZATA DELLA PANCREATITE
ACUTA MODERATA E LA PREDIZIONE DELLE SUE
COMPLICANZE:
VERSO LE FRONTIERE DI UNA MEDICINA PERSONALIZZATA

Tutor
Chiar.mo Prof.
Gianfranco Vallone

Coordinatore
Chiar.mo Prof.
Giovanni Scapagnini

Candidato
Dott. **Marco Di Serafino**
Matr. 172616

Indice

INTRODUZIONE	1
--------------------	---

CAPITOLO I

LA PANCREATITE ACUTA E IL RUOLO DELLA TC

1.1. Definizione ed eziologia	5
1.2. Diagnosi	8
1.3. Classificazione	11
1.4. Diagnostica per immagini: il ruolo della TC	15
1.5. Principali complicanze della pancreatite acuta	20
1.6. Il ruolo delle altre metodiche diagnostiche: ecografia ed RM	24

CAPITOLO II

LA RADIOMICA

2.1 Generalità	30
2.2. Intelligenza artificiale (AI) e machine learning	36
2.3. <i>Software</i> utilizzati	45
2.4. Metodi formali	51
2.5. <i>Model checking</i>	55

CAPITOLO III

OBIETTIVO E STRUTTURA DELLO STUDIO

3.1. <i>Endpoint</i> primario dello studio	58
3.2. Selezione dei pazienti	59
3.3. Acquisizione delle immagini	61
3.4. Segmentazione ed estrazione delle <i>features</i>	63

3.5. Selezione delle <i>features</i>	64
3.6. Utilizzo dei metodi formali.....	74
3.7. Risultati.....	77
3.8. Discussione	80
 IV PROSPETTIVE FUTURE.....	86
 V CONCLUSIONI	96
 Bibliografia	99

INTRODUZIONE

Negli ultimi anni, l'intelligenza artificiale (*Artificial Intelligence* - AI) ha acquisito un ruolo di crescente rilevanza nell'ambito della radiologia, trasformando profondamente le modalità di analisi e interpretazione delle immagini mediche.

La Radiomica in particolare rappresenta una delle applicazioni più innovative dell'AI nella diagnostica per immagini, in quanto permette di estrarre e analizzare caratteristiche quantitative dalle immagini mediche. Questo approccio permette di superare i limiti dell'analisi visiva tradizionale “ad occhio nudo” e di migliorare la performance diagnostica traducendo le immagini in dati numerici utili per identificare marcatori prognostici, predire la risposta a trattamenti, personalizzare le terapie ed altro ancora.

Il presente studio si propone di approfondire il ruolo dell'AI, con particolare riferimento alla Radiomica, nell'ambito della valutazione della pancreatite acuta (PA) nella sua forma moderata. In tale contesto, viene presentata un'analisi innovativa volta a esaminare l'accuratezza di un modello sviluppato mediante l'applicazione dei Metodi Formali per la diagnosi automatizzata delle immagini di Tomografia Computerizzata (TC) di casi di PA moderata.

Quest'ultima rappresenta frequentemente una sfida diagnostica per

il medico radiologo, in quanto non sempre si associa a segni di *imaging* chiaramente identificabili tramite TC, metodica che costituisce il *gold standard* diagnostico di riferimento nei contesti di urgenza [1] [2]. Tale complessità può derivare sia da limitazioni intrinseche alla tecnica stessa, sia da potenziali errori umani, particolarmente significativi in situazioni di urgenza/emergenza. Nonostante ciò, una diagnosi tempestiva riveste un'importanza cruciale, data l'imprevedibilità dell'evoluzione clinica della patologia, che può risolversi spontaneamente oppure degenerare in forme severe di pancreatite, con gravi implicazioni prognostiche per il paziente, fino al rischio di esito fatale [1] [2] [3].

In considerazione di quanto esposto, i promettenti risultati emersi dal presente studio hanno motivato l'estensione dell'applicazione del medesimo modello anche alla previsione dell'evoluzione della PA moderata verso forme severe, come quelle necrotizzanti. Tale ampliamento ha permesso di incrementare il potenziale ambito di utilizzo del modello sviluppato e, conseguentemente, dei Metodi formali, includendo anche applicazioni di natura prognostica.

CAPITOLO I

LA PANCREATITE ACUTA E IL RUOLO DELLA TC

SOMMARIO: 1.1. Definizione ed eziologia – 1.2 Diagnosi –
1.3 Classificazione – 1.4 Diagnostica per immagini: il ruolo della TC –
1.5 Principali complicanze della pancreatite acuta – 1.6 Il ruolo delle altre
metodiche diagnostiche: ecografia ed RM

1.1. Definizione ed eziologia

La PA è un'inflammazione acuta del pancreas, caratterizzata dall'attivazione intra-acinare degli enzimi digestivi pancreatici, che rappresenta il *primum movens* del danno ghiandolare. Questo processo patologico determina un danno acuto d'organo, risultante da un meccanismo di "auto-digestione" tissutale (Figura 1) [1].

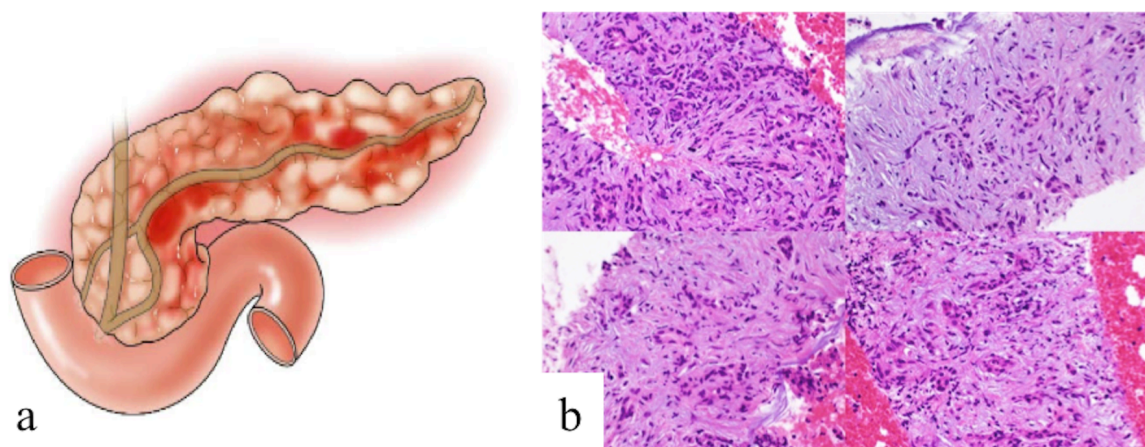


Figura 1: Illustrazione di un processo di PA (fonte: <https://patient.gastro.org/pancreatitis/>) (a) e correlati istologici della distorsione degli acini ghiandolari causata dall'auto-digestione indotta dai succhi pancreatici (b)

L'incidenza della PA è stimata mediamente intorno a 30-45 casi ogni 100.000 abitanti [1].

La PA si distingue per una vasta gamma di manifestazioni cliniche, che possono variare da forme lievi a condizioni più severe, caratterizzate dalla possibile comparsa di complicanze sia locali che sistemiche [2].

Generalmente si presenta con l'insorgenza improvvisa di dolore addominale acuto, riferito principalmente in regione epi-mesogastrica e irradiato al dorso nel 50% dei casi, data la localizzazione retroperitoneale del pancreas. Si possono frequentemente associare altri segni e sintomi quali nausea e/o vomito. Nelle forme lievi, questi sintomi tendono ad essere sfumati e a risolversi entro pochi giorni. Nelle forme più gravi, tuttavia, la PA può progredire fino all'insufficienza multiorgano e al decesso (in oltre il 5% dei casi) [3].

Le principali cause di PA sono rappresentate dalla colelitiasi e dall'abuso di alcool. Meno frequentemente la PA può derivare da disordini metabolici, cause iatrogene, malattie autoimmuni e neoplasie (Tabella 1) [4] [5]. In circa il 20% dei casi l'eziologia resta sconosciuta [5].

- Calcolosi delle vie biliari (circa il 15%)
- Idiopatica (circa il 20%)
- Abuso di alcool
- Disordini metabolici (es. ipertrigliceridemia, ipercalcemia)
- Autoimmune
- Ulcera peptica penetrante
- Trauma
- Iatrogena (es. Colangiopancreatografia retrograda endoscopica - ERCP)
- Neoplasia a lenta crescita
- Malnutrizione
- Pancreatite ereditaria (es. mutazione del gene SPINK 1)
- Infezioni
- Agenti tossici (es. avvelenamento)
- Farmaci (es. steroidi, tetracicline)
- Altro (es. fibrosi cistica, lupus eritematoso sistemico)

Tabella 1. Cause di PA.

1.2. Diagnosi

Prima dell'introduzione diffusa degli esami di *imaging*, nella pratica clinica la valutazione e la stratificazione del rischio nella PA si basavano esclusivamente su sistemi di punteggio clinico, tra cui i punteggi di Ranson, Glasgow e APACHE II (*Acute Physiology and Chronic Health Evaluation*) (Tabella 2) [1] [2].

Tra questi, il punteggio di Ranson è senza dubbio il più noto, nonché uno dei primi ad essere proposti e utilizzati. Tale punteggio si basa sulla valutazione di parametri semplici e di rapida consultazione, distinti in base al momento della rilevazione: all'ammissione (5 parametri, con uno score possibile da 0 a 5) o a 48 ore dall'ammissione (6 parametri, con uno score possibile da 0 a 6) [1] [4] [5].

Il punteggio di Glasgow, derivato da una revisione del sistema di Ranson, è anch'esso largamente impiegato nella pratica clinica. Tuttavia, il punteggio attualmente più utilizzato, grazie alla sua maggiore sensibilità nel prevedere un decorso severo della PA, è il punteggio APACHE II. Quest'ultimo, però, presenta una complessità superiore rispetto ai precedenti, poiché basato su oltre 15 parametri, comportando un maggiore dispendio di risorse e tempo per la sua applicazione [1] [4] [5].

Ranson	Glasgow modificato	APACHE II
Età > 55 anni	Età > 55 anni	Età > 45 anni
Leucociti >16,000/mL	Leucociti >15,000/mL	Leucociti <3,000 o >14,900/mL
Glucosio >200 mg/dL	Glucosio >180 mg/dL	Temperatura rettale <36°C o >38.4°C
LDH >350 IU/mL	LDH >600 IU/L	MAP <70 o >109 mm Hg
AST >250 IU/L	AST >100 IU/L	FC <70 o >109 bpm
Calo dell'Hct >10	Albumina < 3.2 g/dL	FR <12 o >24 bpm
Aumento BUN >5 mg/dL	BUN >96 mg/dL	pH <7.33 o >7.49
Calcio <8 mg/dL	Calcio < 8 g/dL	Na ⁺ <130 o >149 mM
Po ₂ <60 mm Hg	Po ₂ <60 mm Hg	K ⁺ <3.5 o >5.4 mM
Eccesso di basi >4 mEq/L		Po ₂ <70 o >200 mm Hg
Sequestro di liquidi >6 L		Creatinina <0.6 o >1.4 mg/100 mL
		Hct <30% o >45.9%
		GCS

Tabella 2. Confronto tra i tradizionali sistemi di scoring clinico

Un sistema di punteggio sviluppato più recentemente è il BISAP (*Bedside Index of Severity in Acute Pancreatitis*), il cui principale vantaggio risiede nella semplicità d'uso, basandosi su soli cinque parametri: BUN > 8,9 mmol/L, GCS < 15, presenza di SIRS, età > 60 anni e presenza di versamento pleurico. Tuttavia, un limite significativo di questo score è rappresentato dalla sua capacità non ottimale di distinguere tra pazienti con insufficienza d'organo transitoria e

persistente, con il rischio di sovrastimare la gravità del quadro clinico. Tale limite potrebbe compromettere la corretta classificazione tra PA moderata e severa [6].

Sebbene questi punteggi siano utili per stimare la gravità della malattia, risultano insufficienti per prevedere con certezza lo sviluppo di complicanze locali, il rischio di evoluzione in insufficienza d'organo o la mortalità associata alla PA [6].

La classificazione di Atlanta, sviluppata nel 1992, servì per la prima volta a utilizzare termini standardizzati per la PA e le complicanze ad essa associate [7]. Tuttavia, la nomenclatura usata nella classificazione iniziale risultava talvolta confusionaria e la successiva migliore comprensione della fisiopatologia alla base dell'insufficienza d'organo e delle complicanze della PA ha portato a rivedere i suddetti criteri di Atlanta. Attualmente vengono pertanto seguiti i RAC (*Revised Atlanta Criteria*, Criteri di Atlanta Rivisitati) del 2012 [8], secondo cui la diagnosi di PA richiede la presenza di almeno due dei seguenti tre criteri:

1. Dolore epigastrico acuto, intenso e persistente, spesso irradiato al dorso
2. Livello sierico di lipasi o amilasi almeno tre volte superiore al limite massimo della norma

3. Riscontri caratteristici di PA alla TC con mezzo di contrasto (mdc) somministrato per via endovenosa (ev) o, meno frequentemente, alla Risonanza Magnetica (RM) o all'esame ecografico.

Sebbene, in conformità con le suddette linee guida, l'*imaging* possa non essere strettamente necessario per formulare una diagnosi in presenza di criteri clinici e di laboratorio, esso riveste comunque un ruolo fondamentale nella definizione dell'eziopatogenesi e nell'identificazione di eventuali complicanze associate alla PA [7] [8].

1.3. Classificazione

I RAC considerano sia una classificazione temporale che una classificazione in base alla gravità [8].

Classificazione temporale:

- Fase precoce: questa fase, estesa prevalentemente nella prima settimana dall'insorgenza dei sintomi, nei casi di PA moderata-severa è caratterizzata dall'insorgenza di una cascata proinfiammatoria sistemica attivata dall'infiammazione pancreatica, che si manifesta clinicamente come sindrome da risposta infiammatoria sistemica (SIRS), definita dalla presenza di 2 o più criteri diagnostici (Tabella 3). Questa fase può talvolta estendersi anche alla seconda settimana. Il criterio per definire la severità della PA in questa fase precoce è correlato alla presenza e alla durata di

insufficienza d'organo. Quest'ultima viene definita come “*transient organ failure*” se risolta entro le 48 ore o come “*persistent organ failure*” se di durata oltre le 48 ore. Qualora l'insufficienza d'organo includesse più di un sistema d'organo, si definisce un quadro di *multiple organ failure* (MOF) [9].

- Fase tardiva: inizia generalmente nella seconda settimana e può durare mesi, insorgendo pressoché soltanto nei pazienti con PA moderata o grave che sviluppano insufficienza d'organo persistente e/o complicanze locali. In questa fase il ruolo dell'*imaging* è fondamentale per distinguere le caratteristiche morfologiche d'organo e la presenza di eventuali complicanze locali o a distanza; la SIRS infatti può essere seguita da una risposta compensatoria di tipo antinfiammatorio, che può predisporre a complicanze infettive [9].

Temperatura corporea	>38°C o <36°C
Frequenza cardiaca	>90 bpm
Tachipnea	Frequenza respiratoria >20 respiri al minuto o PaCO ₂ <32 mmHg
Leucociti	> 12.000/mm ³ o presenza di >10% di neutrofili immaturi

Tabella 3. Criteri per definire la SIRS

La classificazione della PA per gravità ha un ruolo importante per

definire l'entità ed il timing del trattamento, oltre che l'eventuale necessità di valutazioni specialistiche ulteriori (ad esempio anestesilogiche). Questa classificazione secondo i RAC definisce tre gradi di severità della PA [8]:

- PA lieve: nessuna insufficienza d'organo o complicanze locali/sistemiche.
- PA moderata: insufficienza d'organo transitoria (<48 ore) e/o complicanze locali.
- PA severa: insufficienza d'organo persistente (>48 ore).

Nei casi di PA lieve, il ricorso all'*imaging* diagnostico potrebbe non essere strettamente indispensabile, sebbene un potenziale ruolo possa risiedere nell'identificazione dell'eziopatogenesi della patologia [10].

Nei casi di PA moderata e/o severa, l'*imaging* riveste un ruolo cruciale non solo nell'identificazione della causa della pancreatite, ma anche nella gestione appropriata del paziente in relazione alle complicanze locali o sistemiche [10].

Per la determinazione dell'insufficienza d'organo, il metodo di riferimento approvato è rappresentato dal punteggio di Marshall modificato (Tabella 4) [11]. Tale metodo valuta la funzionalità di tre principali sistemi/apparati: respiratorio, cardiovascolare e renale. A

ciascun sistema/apparato è assegnato un punteggio basato sui risultati ottenuti; un valore pari o superiore a 2 indica la presenza di insufficienza d'organo [11].

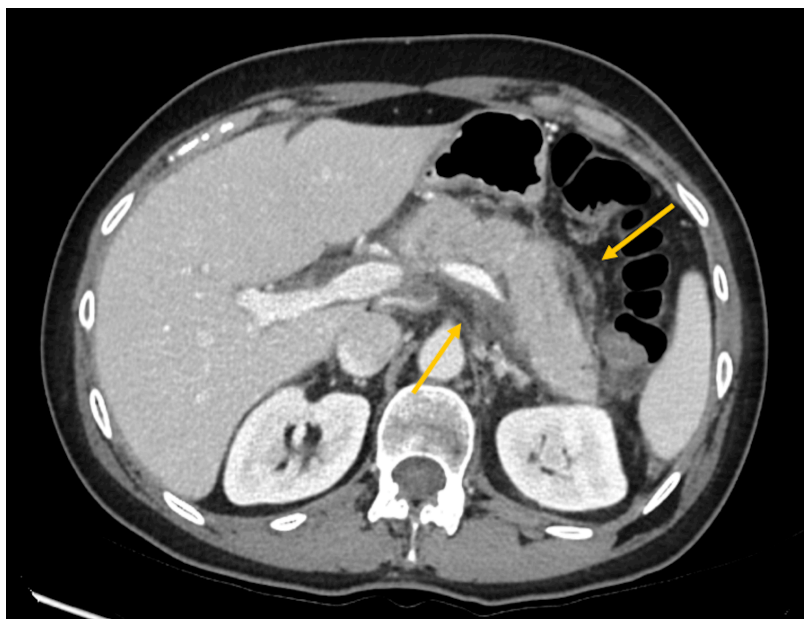
Sistema/ Apparato d'Organo	Score 0	Score 1	Score 2	Score 3	Score 4
Respiratorio (rapporto PaO ₂ /FiO ₂)	> 400	301 – 400	201 – 300	101 – 200	≤ 101
Renale (creatinina sierica)	< 1.4 mg/dL	1.4 – 1.8 mg/dL	1.9 – 3.6 mg/dL	3.6 – 4.9 mg/dL	> 4.9 mg/dL
Cardiovascolare (pressione sistolica, pH)	> 90 mm Hg	< 90 mm Hg, responsiva all'idratazione	< 90 mm Hg, non responsiva all'idratazione	< 90 mm Hg, pH < 7.3	< 90 mm Hg, pH < 7.2

Tabella 4. Punteggio di Marshall modificato.

1.4. Diagnostica per immagini: il ruolo della TC

La TC eseguita con studio multifasico dopo somministrazione di mdc ev riveste un ruolo fondamentale nel percorso clinico-diagnostico della PA. Questa metodica consente di differenziare due principali sottocategorie, definite secondo i criteri morfologici stabiliti dai RAC [12] [13], ovvero:

- PA interstiziale/edematosa: caratterizzata da diffuso (più raramente localizzato) ingrossamento del pancreas per la presenza di edema infiammatorio. L'*enhancement* post-contrastografico della ghiandola pancreatica risulta sostanzialmente conservato o solo minimamente ridotto a causa dei fenomeni edemigeni. Vi è solitamente un'imbibizione edemigena e/o un moderato "*stranding*" del cellulare adiposo peripancreatico. Può riscontrarsi anche una quota fluida peripancreatica (Figura 2).



***Figura 2.** Caso di PA edematosa-interstiziale evidenziato con esame TC con mdc ev in fase venosa di acquisizione di studio. Le frecce indicano la presenza di un moderato fat-stranding peripancreatico con esile falda fluida peripancreatica.*

- PA necrotizzante: circa il 5-10% dei pazienti con PA può sviluppare una necrosi del parenchima pancreatico, del cellulare adiposo peripancreatico o, più comunemente, di entrambi. Questo tipo di evento, tuttavia, è riscontrabile alla TC solo alcuni giorni dopo l'insorgenza della pancreatite, per cui una TC eseguita troppo precocemente dopo l'insorgenza dei sintomi può sottostimare l'eventuale evoluzione in pancreatite necrotizzante. La necrosi pancreatico e peri-pancreatica è rappresentata da un'irregolare area, più o meno estesa, di scarsa impregnazione post-contrastografica (Figura 3).

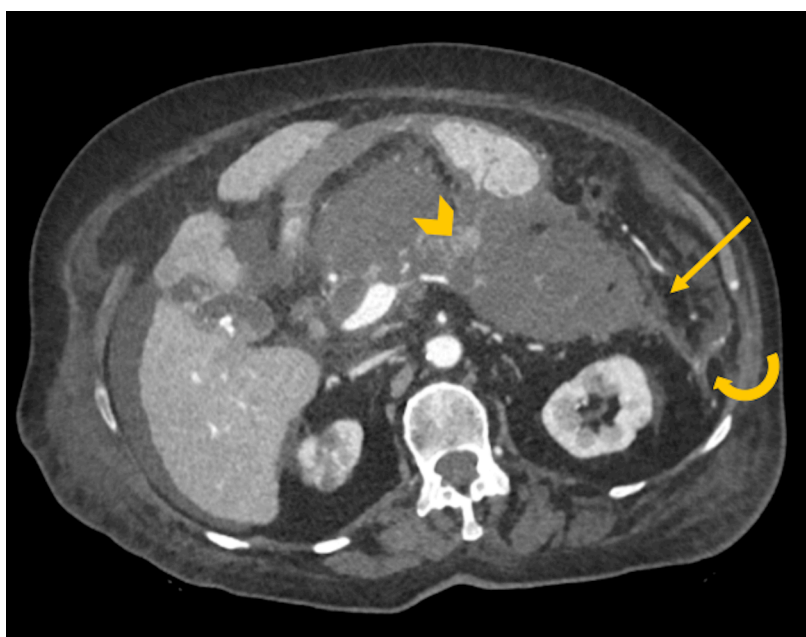


Figura 3. Caso di PA necrotico-emorragica severa evidenziato mediante TC con mdc ev in fase venosa di acquisizione di studio. La scansione mostra un completo sovvertimento della regione pancreatica, con estese aree prive di enhancement post-contrastografico (freccia) e una minima porzione pancreatica residua localizzata nel corpo (punta di freccia). Si osserva inoltre la presenza di una falda di liquido libero addominale (freccia curva).

Le aree di necrosi pancreatica o peri-pancreatica possono rimanere sterili oppure essere complicate da un'infezione, configurando un quadro di pancreatite necrotizzante infetta. La presenza di sovra-infezione può essere sospettata in base all'evidenza radiologica di aria all'interno delle aree necrotiche [12] [13].

Come già accennato, è fondamentale conoscere con precisione il tempo trascorso dall'insorgenza dei sintomi, al fine di pianificare adeguatamente il corretto *timing* dell'esame TC. In particolare, nelle prime 72 ore dall'inizio dei sintomi, i reperti radiologici possono essere poco evidenti, aumentando il rischio di un'errata valutazione della gravità della PA [8].

Nel 1990 è stato introdotto un indice di gravità basato sui reperti tomografici, noto come *Computed Tomography Severity Index* (CTSI) [12] [14]. Questo indice combina i criteri di Balthazar (attribuendo un punteggio da 0 a 4 punti) con un punteggio aggiuntivo basato sull'estensione della necrosi pancreatica (da 0 a 6 punti), determinando un punteggio complessivo che varia da 0 a 10 (Tabella 5). Un punteggio più elevato è associato a un incremento proporzionale della morbidità e della mortalità [12] [13] [14].

Grado di pancreatite (Balthazar)	Punti
Pancreas normale	0
Pancreas ingrossato	1
Infiammazione peripancreatica	2
1 raccolta fluida peripancreatica	3
>1 raccolte fluide peripancreatiche	4
Necrosi pancreatica	Punti
Assente	0
< 30%	2
30 - 50%	4
> 50%	6

Tabella 5. CTSI

Nel 2004, Mortelet et al. hanno suggerito una versione modificata del CTSI (mCTSI) [15], semplificando sia la valutazione dell'infiammazione pancreatica che l'estensione della necrosi pancreatica (soltanto < 30% o > 30%), ma aggiungendo alcuni ulteriori fattori quali versamento pleurico, ascite, complicanze vascolari, parenchimali pancreatiche o

gastrointestinali (Tabella 6).

Infiammazione (punteggio)	Necrosi (punteggio)	Complicanze extra- pancreatiche (una o più, 2 punti)
Pancreas normale (0)	Assente (0)	Versamento pleurico, ascite, complicanze vascolari, complicanze parenchimali, coinvolgimento del tratto gastrointestinale
Anomalie del pancreas con o senza alterazioni infiammatorie del grasso peripancreatico (2)	< 30% (2)	Versamento pleurico, ascite, complicanze vascolari, complicanze parenchimali, coinvolgimento del tratto gastrointestinale
Raccolte fluide pancreatiche o peripancreatiche o necrosi del grasso peripancreatico (4)	> 30% (4)	Versamento pleurico, ascite, complicanze vascolari, complicanze parenchimali, coinvolgimento del tratto gastrointestinale

Tabella 6. mCTSI

Sono stati condotti numerosi studi comparativi sui vari sistemi di punteggio, in particolare tra i RAC, il CTSI e la sua versione modificata mCTSI. Complessivamente, si osserva una buona concordanza tra questi punteggi; tuttavia, secondo alcune analisi, il mCTSI si è dimostrato più

sensibile, ma meno specifico rispetto al CTSI nel distinguere la pancreatite acuta moderata da quella severa [16].

Altri studi, invece, hanno evidenziato che il mCTSI risulta più accurato, più semplice da calcolare e caratterizzato da una minore variabilità inter-osservatore rispetto al CTSI, rendendolo potenzialmente più efficace nella gestione dei pazienti con PA [17].

1.5. Principali complicanze della pancreatite acuta

Tra le complicanze della PA rientrano sia complicanze locali che complicanze a distanza. Per quanto riguarda le complicanze locali, le più comuni sono le raccolte fluide pancreatiche e peripancreatiche.

I RAC del 2012 distinguono 4 tipi di raccolte in base al contenuto, al grado di “organizzazione” e al tempo intercorso dall’insorgenza della PA [8] [12] [13]. La nuova classificazione di Atlanta, in particolare, abbandona i termini fuorvianti utilizzati in ambito clinico, come "pseudocisti acuta" o "ascesso," per descrivere qualsiasi raccolta peri-pancreatica in presenza di infiammazione PA. Inoltre, introduce una distinzione significativa tra raccolte caratterizzate da contenuto fluido omogeneo, tipiche della PA interstiziale edematosa, e raccolte dall'aspetto eterogeneo, contenenti detriti necrotici, tipiche della PA necrotizzante [8].

I RAC includono quindi nuove definizioni che descrivono più accuratamente i vari tipi di raccolte, introducendo in particolare termini come “APFC” (*Acute Peripancreatic Fluid Collection*, Raccolta fluida peripancreatica acuta), “ANC” (*Acute Necrotic Collection*, Raccolta necrotica acuta), “Pseudocisti” e “WON” (*walled-off necrosis*, Necrosi murata) (Figura 4), in base alla durata (meno di 4 settimane o più di 4 settimane dall'insorgenza del dolore), al grado di organizzazione e alla presenza o assenza di necrosi all'imaging (Tabella 7) [7].

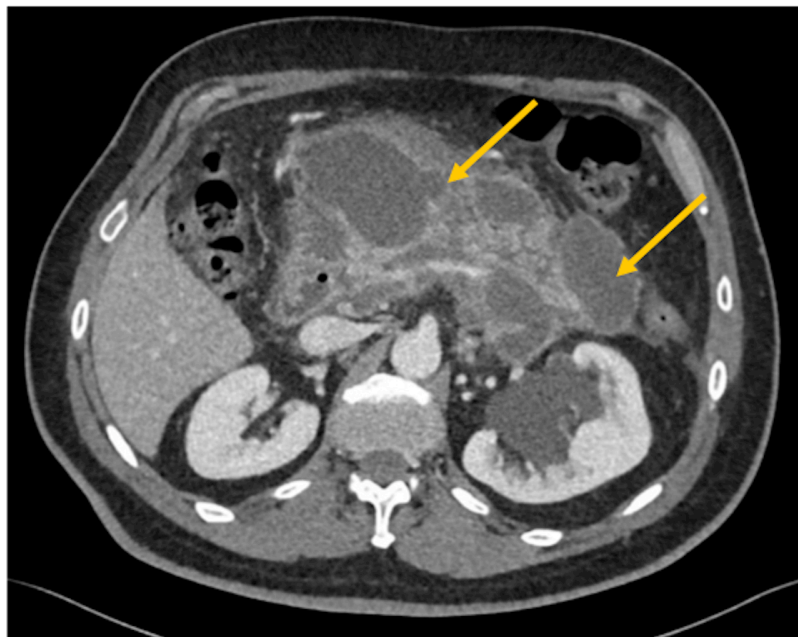


Figura 4. Evoluzione di una PA necrotico-emorragica con evidenza di estese raccolte necrotiche intra- e peri-ghiandolari (freccie).

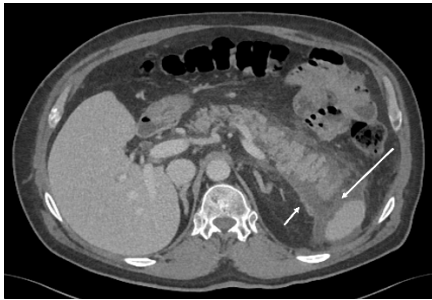
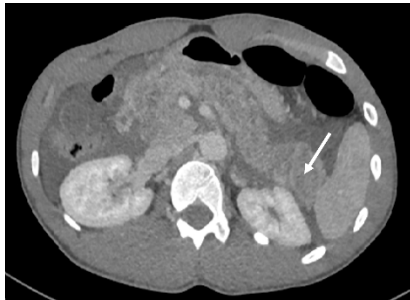
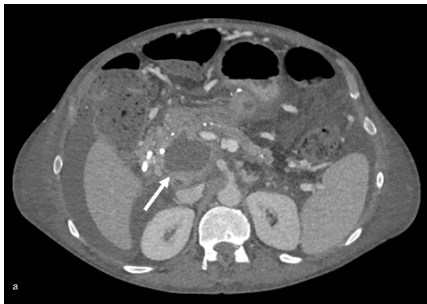
	Pancreatite interstiziale -edematosa	Pancreatite necrotizzante
< 4 settimane	APFC: Raccolta fluida omogenea adiacente al pancreas senza una parete ben riconoscibile (freccie) 	ANC: Raccolta necrotica intra e/o extra-pancreatica senza una parete ben definita (freccia) 
> 4 settimane	Pseudocisti: Raccolta fluida capsulata ben definita, solitamente extra- pancreatica con possibili detriti proteinaceo-ematici (freccia) 	WON: Raccolta necrotica intra e/o extra-pancreatica con una parete ben definita (freccie) 

Tabella 7. Classificazione delle raccolte peri-pancreatiche

Ciascuna delle raccolte sovra-descritte può eventualmente sviluppare una sovrainfezione batterica, specie nel caso delle raccolte necrotiche. La sovrainfezione è abbastanza rara nella prima settimana ma può essere sospettata alla TC dalla presenza di bolle d'aria nelle raccolte, spesso secondarie alla colonizzazione da parte di bacilli gram-negativi enterici [18].

Le raccolte possono però, più raramente, essere anche la conseguenza di una fistola pancreatico-enterica derivante da raccolte necrotiche capaci di erodere la parete intestinale (principalmente a livello di colon o duodeno) [19].

Altre complicanze della PA includono:

- Effetto massa sugli organi adiacenti: in particolare nei casi di pancreatiti necrotico-emorragiche complicate da voluminose raccolte, con conseguente possibile insorgenza di ostruzioni meccaniche gastrointestinali, idronefrosi o ostruzione delle vie biliari [19].
- Complicanze a carico del dotto di Wirsung: come il "dotto pancreatico disconnesso" (causato dalla presenza di una raccolta localizzata a livello del collo/corpo pancreatico o dalla necrosi parenchimale stessa) o stenosi del dotto pancreatico, quest'ultima più frequente come complicanza tardiva [20].
- Complicanze vascolari: associate ad elevata morbilità e mortalità, tra cui la rottura vascolare con emorragia o la formazione di pseudoaneurismi (Figura 5) [21].

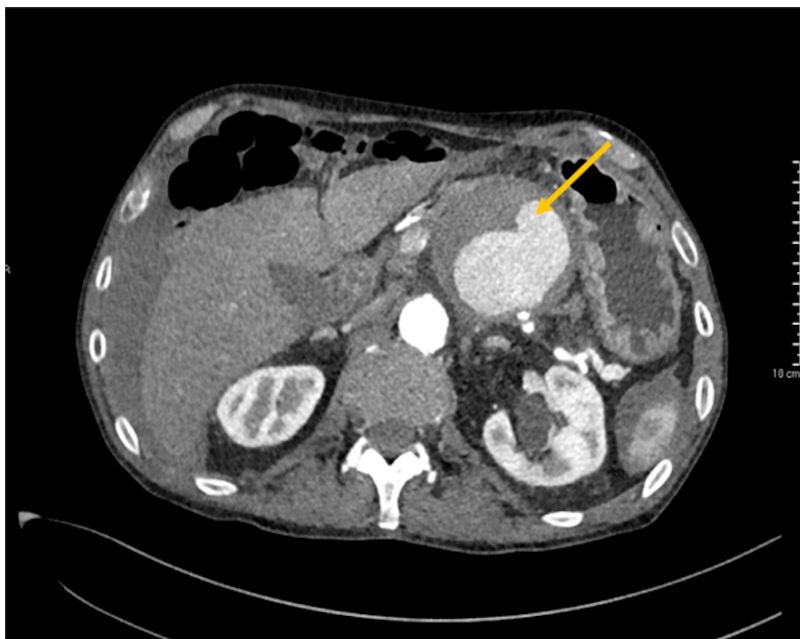


Figura 5. PA necrotico-emorragica complicata da voluminoso pseudoaneurisma (freccia) dell'arteria splenica.

1.6. Il ruolo delle altre metodiche diagnostiche: ecografia ed RM

L'ecografia addominale rappresenta una delle metodiche di imaging di prima scelta nei pazienti con sospetta PA. Grazie alla sua ampia disponibilità, rapidità di esecuzione e sicurezza, risulta particolarmente adatta per una valutazione iniziale. Tra i segni caratteristici osservabili figurano l'aumento delle dimensioni del pancreas, la presenza di contorni irregolari e un'ecostruttura eterogenea riconducibile al processo infiammatorio. In alcune circostanze, l'ecografia può evidenziare raccolte fluide peripancreatiche o edema, reperti tipici delle fasi iniziali della malattia (Figura 6) [22].

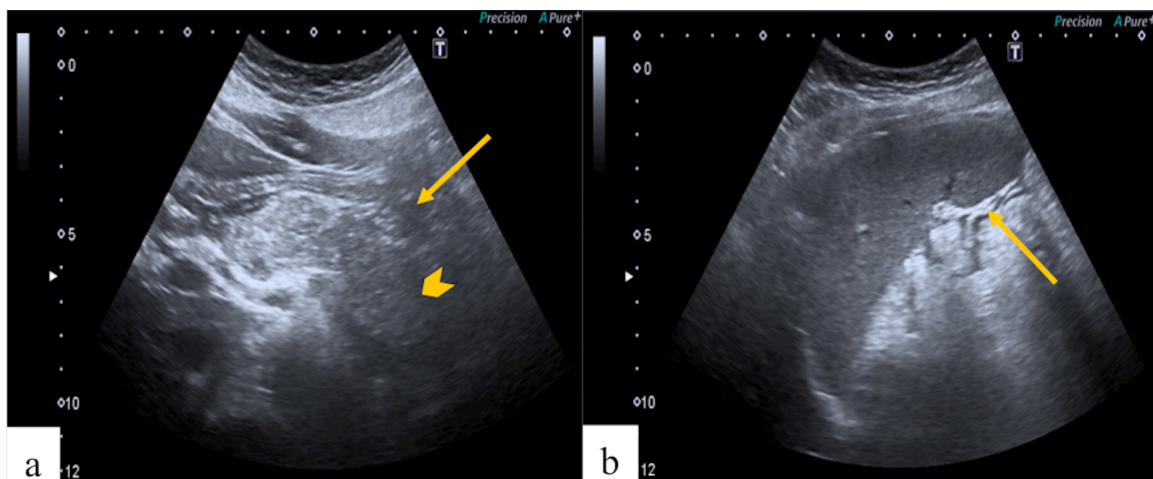


Figura 6. a) PA interstiziale-edematosa con evidenza ecografica di tumefazione del corpo-coda pancreatico (punta di freccia) con presenza di falda fluida peri-pancreatica (freccia); b) estensione della falda fluida in sede peri-splenica (freccia).

L'ecografia è uno strumento diagnostico particolarmente utile per l'identificazione dei calcoli biliari, che rappresentano una delle principali cause di PA. Questa metodica consente di rilevare la presenza di calcoli sia nella colecisti sia nelle vie biliari (Figura 7) [22].

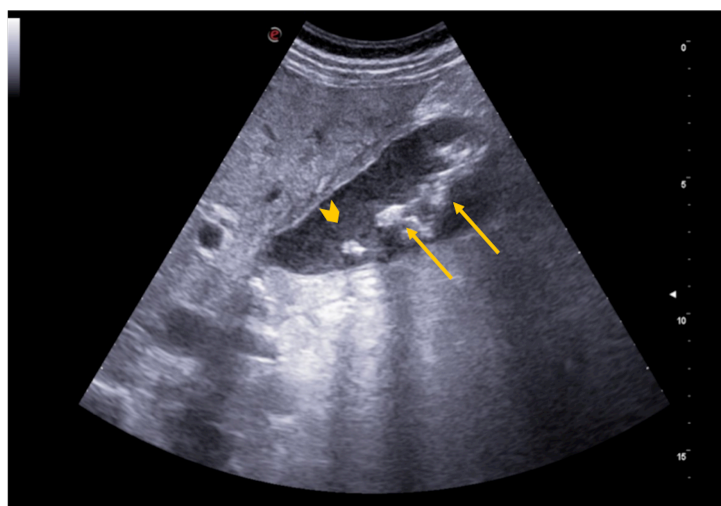


Figura 7. Evidenza ecografica di sludge biliare intracolecistico in sede declive (punta di freccia) e di multipli aggregati litiasici (freccie) in un paziente affetto da PA interstiziale-edematosa.

Tuttavia, l'ecografia presenta limitazioni nella valutazione del tratto terminale del coledoco, per il quale risulta più indicato l'utilizzo della RM con sequenze colangio. Una diagnosi tempestiva di pancreatite biliare consente una gestione mirata, che può includere procedure endoscopiche per la rimozione dei calcoli [22] [23].

Durante la fase di remissione, l'ecografia si rivela uno strumento utile per monitorare la risoluzione delle raccolte fluide e per rilevare eventuali complicanze, quali la formazione di pseudocisti o infezioni. Tuttavia, l'affidabilità di questa metodica può essere compromessa dalla presenza di gas intestinale o obesità, fattori che possono limitare significativamente la qualità della visualizzazione [22] [23].

La RM offre una rappresentazione altamente dettagliata del pancreas e delle strutture circostanti. Questa metodica è particolarmente utile per caratterizzare la gravità della pancreatite, distinguendo tra edema, infiammazione e necrosi. Sequenze T2-pesate evidenziano raccolte fluide e alterazioni infiammatorie, mentre sequenze T1-pesate possono identificare infiltrati adiposi o aree di sanguinamento [23].

Le sequenze colangio-RM consentono una accurata visualizzazione delle vie biliari e dei dotti pancreatici. Questa procedura è utile soprattutto nei casi di sospetta pancreatite di origine biliare, stenosi del dotto pancreatico o anomalie congenite (Figura 8) [24]. La capacità di identificare calcoli o

ostruzioni senza l'uso di mezzo di contrasto iodato la rende particolarmente indicata nei pazienti con insufficienza renale o allergie al mdc [23] [24].

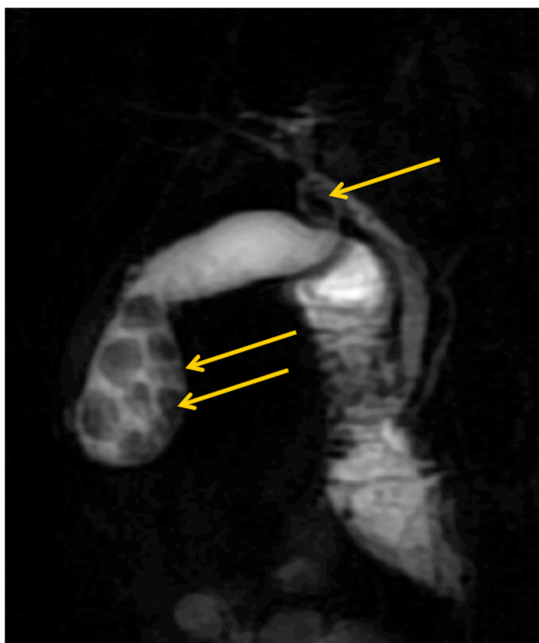


Figura 8. Risonanza Magnetica con sequenze colangio evidenzia la presenza di aggregati litiasici all'interno del lume colecistico e nel dotto cistico, associati a una dilatazione consensuale delle vie biliari, in un paziente affetto da PA interstiziale-edematosa.

In più, le complicanze tardive, come pseudocisti e WON, richiedono una caratterizzazione dettagliata per pianificare eventuali interventi. La RM consente infatti di distinguere tra raccolte fluide semplici e raccolte complesse a contenuto proteinaceo, necrotico-emorragico (Figura 9). Questo livello di dettaglio è fondamentale per decidere se procedere con drenaggi percutanei o approcci chirurgici [25].

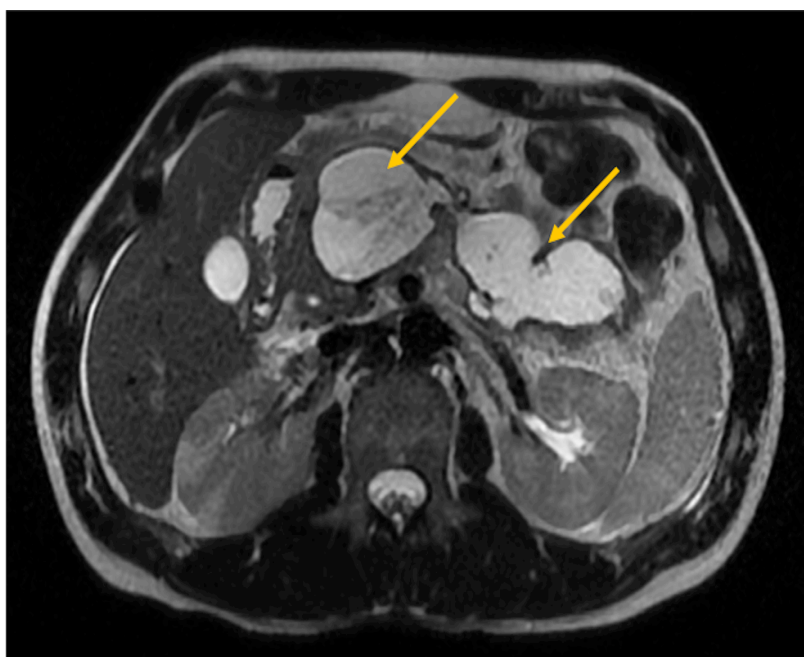


Figura 9. Risonanza Magnetica dell'addome superiore, sequenza T2-pesata: evidenza di raccolte necrotiche organizzate -WON (frecce), in un paziente con pregressa PA necrotico-emorragica.

L'ecografia e la RM svolgono ruoli distinti ma complementari nella gestione della PA. L'ecografia, grazie alla sua rapidità e ampia disponibilità, rappresenta spesso la metodica di prima linea per una diagnosi tempestiva e per l'identificazione di cause di origine biliare. La RM, invece, offre un livello di dettaglio superiore ed è preferita nei casi più complessi o per il monitoraggio delle complicanze. Inoltre, è importante sottolineare che al pari dell'ecografia, anche la RM, non prevede l'utilizzo di radiazioni ionizzanti, rendendola una scelta particolarmente indicata nei pazienti giovani o in quelli che necessitano di controlli ripetuti [25]. La selezione della metodica diagnostica è influenzata dalla disponibilità della struttura, dalle condizioni cliniche del paziente e dalla fase della malattia [25].

CAPITOLO II

LA RADIOMICA

SOMMARIO: 2.1. Generalità – 2.2 Intelligenza artificiale (AI) e machine learning – 2.3 Software utilizzati – 2.4 Metodi formali – 2.5 Tecnica del model checking

2.1. Generalità

La Radiomica rappresenta un'innovativa frontiera della medicina, focalizzata sull'estrazione di dati quantitativi dalle immagini radiologiche e sul loro utilizzo per sviluppare sistemi a supporto delle decisioni cliniche [26]. Questo approccio risulta molto promettente in quanto permette di identificare informazioni che, altrimenti, potrebbero non essere percepite dall'occhio umano.

I dati quantitativi ottenuti dalle immagini possono essere integrati in modelli multidisciplinari sia morfologici che predittivi, utili per la gestione clinica del paziente. Questi modelli si avvalgono anche di tecniche di *Machine Learning* e AI [26]. Le caratteristiche morfologiche e predittive si dividono in gruppi distinti: dimensioni, proprietà morfologiche e descrittori che analizzano le relazioni tra i *pixel* o *voxel* presenti nell'immagine. Le definizioni matematiche di queste relazioni, chiamate "descrittori", non dipendono dalla modalità di *imaging*, ma i loro valori sono influenzati dalle impostazioni della macchina radiologica. In genere, questi descrittori mostrano come i valori numerici presenti all'interno delle immagini siano distribuiti spazialmente (e non) all'interno del tessuto studiato [27][28].

Considerando che a ogni intensità di grigio in un'immagine digitale corrisponde un valore numerico specifico, la distribuzione spaziale di

questi valori, all'interno delle strutture analizzate, è influenzata da diversi fattori legati alla metodica di *imaging*. Questo tipo di analisi viene denominato “*texture analysis*” [29][30].

Nelle immagini TC, per esempio, l'intensità riflette la radiopacità dei tessuti, e la distribuzione spaziale di questo parametro in una lesione tumorale può risentire della presenza di necrosi o calcificazioni, elementi spesso correlati all'aggressività della malattia. Simili analisi possono essere condotte anche su immagini ecografiche [31], di RM [32][33], Tomografia ad Emissione di Positroni (PET) [34] e Tomoscintigrafia (SPECT) [35], applicando la medesima base matematica.

Le tecniche di radiodiagnostica offrono un'elevata risoluzione spaziale, che consente di analizzare molti *voxel* o *pixel* e di ottenere descrittori statistici numericamente robusti. Le metodiche medico-nucleari, invece, pur avendo una risoluzione spaziale più bassa, vantano una migliore risoluzione di contrasto, ideale per studiare i processi biologici nelle lesioni sospette. Tuttavia, questa limitazione può rappresentare un problema nello studio di lesioni di dimensioni ridotte [29] [30].

In oncologia, i descrittori ottenuti con queste tecniche sono stati frequentemente correlati con l'analisi genetica delle lesioni, evidenziando una correlazione tra le diverse strutture tumorali e le mutazioni genetiche

associate alle cellule. Questo approccio prende il nome di radiogenomica [29] [30] [36].

La medicina di precisione, considerata uno degli obiettivi principali della medicina contemporanea, si propone di personalizzare le terapie in funzione delle specifiche caratteristiche del paziente e della patologia da cui è affetto. In questo contesto, la Radiomica si sta affermando rapidamente come una tecnologia di fondamentale importanza, rappresentando uno dei settori di ricerca più promettenti. Tra le applicazioni più rilevanti della Radiomica figura l'oncologia, dove, attraverso l'*imaging* diagnostico, è possibile raccogliere informazioni dettagliate e caratterizzare i tessuti in maniera non invasiva, riducendo così la necessità di ricorrere a biopsie o interventi chirurgici tradizionali [35] [36] [37].

L'enorme patrimonio di dati che la Radiomica ci permette di ottenere ed analizzare, quantitativo di dati che sarebbe impossibile elaborare con la sola osservazione visiva anche da parte dell'operatore più esperto, permette di definire numerose caratteristiche della "lesione" d'interesse oltre che dell'ambiente/tessuto circostante, caratteristiche relative ad esempio alla forma, al volume e alla struttura [35] [36] [37]. Con queste metodiche risulta quindi possibile studiare non solo le caratteristiche stesse dei dati analizzati ma anche l'eventuale correlazione fra i dati

ottenuti dalle immagini e le caratteristiche molecolari o genomiche della patologia, con lo scopo finale di poter estrapolare direttamente dall'analisi delle immagini indicazioni sull'aggressività della malattia, ad esempio, o sulle terapie più adeguate e l'eventuale risposta ad esse. Per questo motivo si ha la necessità di realizzare un "classificatore" [35] [36] [37], la cui effettiva capacità di effettuare diagnosi più o meno corrette viene definita sulla base di quattro metriche principali:

- *True Positive* (TP): indica i pazienti malati classificati correttamente come tali; la "T" di "*True*" è il *Gold standard* (in questo caso il medico radiologo che è a conoscenza della verità), mentre la "P" di "*Positive*" è il risultato del classificatore;
- *True Negative* (TN): indica i pazienti sani correttamente classificati come sani;
- *False Positive* (FP): indica i pazienti sani erroneamente classificati come malati (in tale caso si ha quindi una classificazione errata);
- *False Negative* (FN): indica i pazienti malati erroneamente classificati come sani (anche in questo caso il risultato del classificatore è errato).

L'errore di classificazione può essere calcolato utilizzando la cosiddetta *Confusion Matrix* (Figura 10), una matrice composta da due

righe e due colonne, nella quale vengono registrati e sommati tutti gli elementi derivanti dall'analisi [29] [30] [36].

A partire dalle metriche sopra descritte, è possibile derivare quattro indici utili a definire in modo più dettagliato le caratteristiche e l'accuratezza del classificatore:

- *Accuracy*: rappresenta il numero di pattern correttamente classificati. In genere si preferisce esprimerlo in percentuale, indicando così la percentuale di classificazioni corrette. Tanto più tale valore è vicino ad 1, tanto più efficiente sarà il classificatore.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- *Precision*: è un indice di correttezza che valuta i pazienti correttamente classificati come positivi sul numero complessivo di pazienti positivi.

$$Precision = \frac{TP}{TP + FP}$$

- *Recall*: è una misura di completezza, indicata il più delle volte con il termine di *Sensitivity*.

$$Recall = \frac{TP}{TP + FN}$$

- *Specificity*: è un parametro che valuta il numero di negativi che sono correttamente classificati come tali; viene indicato anche come Vero Tasso Negativo.

$$Specificity = \frac{TN}{TN + FP}$$

		Valori effettivi		
		+	-	
Val. previsti	+	True positive (TP)	False positive (FP)	Precision = TP / (TP+FP)
	-	False negative (FN)	True Negative (TN)	
		Recall = TP/(TP+FN)		Accuracy = (TP+TN) / (TP+FP+TN+FN)

Figura 10. Un esempio di matrice di confusione

2.2. Intelligenza artificiale (AI) e Machine Learning

L'AI costituisce la base per simulare i processi di intelligenza umana attraverso la creazione e l'applicazione di algoritmi integrati in un ambiente di calcolo dinamico. Ciò viene effettuato o attraverso l'elaborazione di modelli ideali o, concretamente, con la creazione di macchine che utilizzano principalmente elaboratori elettronici. Propriamente, l'AI è una disciplina appartenente al campo dell'informatica: essa studia i fondamenti teorici, le metodologie e le tecniche per progettare sistemi *hardware* e programmi *software* [29] [30] [36].

La “data di nascita” dell'AI viene fissata nell'anno 1956, in cui per la prima volta si parlò di Sistema Intelligente nel contesto di un convegno in America, come precursore, a posteriori, dell'attuale AI. In tale contesto fu infatti presentato un programma, sviluppato da due informatici, capace di dimostrare “autonomamente” alcuni teoremi matematici partendo da informazioni prestabilite e fornite dai programmatori [29] [30] [36]. A partire da quel momento, furono effettuate numerose scoperte che consentirono lo sviluppo di programmi in grado di dimostrare teoremi di crescente complessità. In questo contesto nacque Lisp (*List Processor*), il primo linguaggio di programmazione dedicato, che per molti anni costituì la base

fondamentale per i software di AI.

Una sfida particolarmente interessante, per quanto ardua, del periodo di evoluzione e miglioramento progressivo dell'AI fu quella mirata a implementarne il “linguaggio appropriato” per giungere alla soluzione conclusiva [29] [30] [36]. In altre parole, per elaborare soluzioni più complesse, il software o la macchina necessitavano di “ragionare” e prendere decisioni opportune anche mediante l'evoluzione parallela di un appropriato linguaggio e di percorsi semantici adatti [29] [30] [36].

Un contributo significativo alla ricerca sull'AI deriva, in misura rilevante, non tanto dal campo informatico in senso stretto, quanto da quello biologico. L'AI rappresenta infatti uno strumento potente, già parzialmente integrato nella medicina moderna. Le sue applicazioni spaziano dalla diagnostica alla chirurgia, dallo sviluppo di farmaci alla riabilitazione, ambiti destinati senza dubbio a un'ulteriore espansione [29] [30] [36].

In particolare, nel settore medico, l'AI si propone di perseguire due finalità fondamentali:

- Cercare di implementare le “capacità” diagnostiche e terapeutiche di un operatore sanitario ed in generale di alleviare il carico di lavoro sul singolo;

- Supportare la gestione più “globale” di un sistema sanitario al fine di migliorarne le prestazioni.

Nel campo della radiologia, l'AI riveste un ruolo di particolare rilievo, poiché consente di identificare dettagli e caratteristiche che sfuggono all'osservazione diretta dell'occhio umano.

Un presupposto fondamentale per comprendere appieno l'utilizzo e l'importanza dell'AI in radiologia è il concetto che l'immagine radiologica, così come qualsiasi immagine digitale, può essere considerata una vera e propria “banca dati”, ovvero un insieme strutturato di dati e informazioni [28]. Quando un medico radiologo osserva e analizza un esame radiologico, formula una diagnosi e redige un referto basandosi sulla propria esperienza e conoscenza, confrontando le immagini con ciò che ha acquisito nel corso della propria carriera professionale.

L'AI esegue sostanzialmente un processo analogo, ma con una capacità di analisi e confronto di immagini significativamente superiore [29] [30] [36]. Ciò avviene attraverso un processo di apprendimento dei dati contenuti nelle immagini, noto come *Machine Learning*. Quest'ultimo rappresenta un sottoinsieme dell'AI, che permette a una macchina di acquisire conoscenze e migliorare le proprie prestazioni senza la necessità di una programmazione esplicita per ciascun compito

specifico [29] [30] [36].

L'“addestramento” del sistema in questo contesto richiede l'utilizzo di ingenti quantità di dati, unitamente a un algoritmo efficiente, in grado di individuare schemi e correlazioni per formulare decisioni e previsioni ottimali basate sulle analisi effettuate. Questo processo di apprendimento si fonda sui dati, intesi come elementi informativi rappresentati da simboli che necessitano di essere elaborati, e sui pattern, elemento cruciale del *Machine Learning*. In tale ambito, il comportamento degli algoritmi non è pre-programmato, ma viene appreso direttamente dagli stessi. È comune, infatti, utilizzare il termine "pattern" per riferirsi ai dati analizzati [29] [30] [36].

Esistono diversi tipi di pattern:

- Numerici: valori associati a caratteristiche misurabili o conteggi, tipicamente continui ma soggetti ad ordinamento e rappresentabili come vettori numerici nello spazio multidimensionale. L'estrazione di caratteristiche da segnali produce vettori numerici detti *feature vectors*.
- Categorici: valori associati a caratteristiche qualitative e alla presenza/assenza di una caratteristica; non sono “semanticamente” mappabili in valori numerici.

- Sequenze: pattern sequenziali con relazioni spaziali o temporali, spesso a lunghezza variabile. Larga importanza è data alla posizione nella sequenza in relazioni a predecessori e successori.

La complessità del *Machine Learning* ha reso necessaria l'individuazione di tre diverse tipologie di apprendimento, distinte in base alle specifiche esigenze poste al sistema: apprendimento supervisionato, apprendimento non supervisionato e apprendimento semi-supervisionato (o apprendimento per rinforzo) [29] [30].

Nell'apprendimento supervisionato, in particolare, alla macchina vengono forniti degli esempi di obiettivi da raggiungere, per "indirizzarla" mostrando le relazioni tra *input*, *output* e risultato. Dall'insieme dei dati mostrati, la macchina deve essere in grado di estrapolare una regola generale che possa permettere, ogni volta che venga stimolata con un determinato *input*, di scegliere l'*output* corretto per il raggiungimento dell'obiettivo.

Nel caso di apprendimento non supervisionato, invece, la macchina non viene indirizzata in alcun modo, dunque è come se "imparasse" soltanto dagli errori commessi.

Infine, nell'apprendimento per rinforzo la macchina si trova ad interagire con un ambiente in cui le caratteristiche sono variabili, dunque un ambiente dinamico, nel quale la macchina dovrà muoversi, in maniera

“casuale” e non indirizzata preventivamente, verso un obiettivo, potendo avere la possibilità, soltanto a prova conclusa, di sapere se è riuscita o meno ad adempiere al compito ad essa assegnato.

Nel contesto radiologico, il *Machine Learning* viene impiegato sottoponendo al sistema una serie di immagini, previamente classificate come sane o patologiche, al fine di "insegnare" al sistema a distinguerle in base alle informazioni contenute in esse. Il *software*, progressivamente, immagazzina le informazioni acquisite durante il processo di apprendimento e, una volta completato tale apprendimento, sarà in grado di applicare la propria "conoscenza" per analizzare e classificare nuove immagini che gli verranno sottoposte [29] [30] [36].

Attualmente esistono diversi dispositivi basati su tecniche di AI nel campo del processamento e dell'analisi di immagini radiologiche, approvati nell'Unione Europea e negli USA, che fungono da supporto alla diagnosi, attraverso la rivelazione di anomalie, con informazioni quantitative mostrate generalmente nei sistemi *Computer-Aided Design* (CAD) per agevolare il medico radiologo nell'individuare eventuali reperti patologici [29] [30] [35] [36].

La costante ricerca di diagnosi sempre più precise e mirate e di minore invasività per il paziente ha ulteriormente stimolato, anche in modo caotico, sviluppi e proposte di utilizzo di nuove tecnologie nel

campo radiologico. Tra queste l'AI, negli ultimi anni in rapidissima espansione, rappresenta indubbiamente la tecnologia più esplorata. La ragione di questo interesse è dovuta sia alle numerose implementazioni disponibili, consolidate e certificate, sia alla relativa facilità di migliorare o riadattare strumenti software in tempi relativamente brevi. Gli strumenti basati sull'apprendimento automatico (*Data-driven*) dovrebbero essere efficaci quando non è ancora disponibile una modellizzazione quantitativa del problema, a causa della sua complessità o novità, ma è già presente una quantità di dati statisticamente significativa.

Nella pratica un processo standard di analisi delle immagini in Radiomica può essere così sintetizzato [29] [30]:

- Acquisizione delle immagini;
- Identificazione del volume di interesse (ROI, *Region of Interest*), da cui estrarre i dati;
- Segmentazione del volume di interesse mediante strumenti grafici;
- Estrazione dei dati dalle immagini digitali;
- Creazione di modelli utili per la classificazione.

A queste fasi possono esserne aggiunte ulteriori per affinare l'analisi. Ad esempio, dopo l'acquisizione dell'immagine, può essere

eseguito un miglioramento della qualità della stessa (in termini spaziali con il *Resampling*); può anche essere eseguito il cosiddetto filtraggio, ovvero l'implementazione di filtri sull'immagine al fine di ridurre il rumore o esaltare alcune caratteristiche. Inoltre, può essere effettuata una selezione dei dati per creare modelli predittivi di successo al fine di eliminare la ridondanza delle informazioni e accelerare le prestazioni [29] [30].

I dati estratti in questo processo sono definiti “*features* radiomiche” e ne esistono due tipi: semantiche e agnostiche [29] [30].

- Le *features* semantiche sono quelle comunemente utilizzate nel lessico radiologico (dimensione, forma, posizione, viscosità, spiculatoria dei margini, necrosi ecc...) e sono quindi facilmente collegate alle evidenze radiologiche.
- Le *features* agnostiche sono costituite da dati calcolati utilizzando formule matematiche basate sulle informazioni nell'immagine, come l'istogramma o le matrici costruite in base alla relazione spaziale dei livelli di grigio nei *pixel*. Tra queste vi sono la *texture* di Haralick [37], Laws [38], Wavelets [39], trasformate di Laplace, funzionali di Minkowski e frattali [38] [39].

Nella categoria delle *features* agnostiche, si possono trovare caratteristiche di primo, secondo o ordine superiore.

Le caratteristiche di primo ordine si basano sull'istogramma, ossia i livelli di grigio nell'immagine senza tener conto della loro relazione spaziale. Ad esempio, non sono in grado di distinguere tra un motivo a scacchiera in bianco e nero e un'immagine divisa con un lato nero e l'altro bianco (Figura 11).

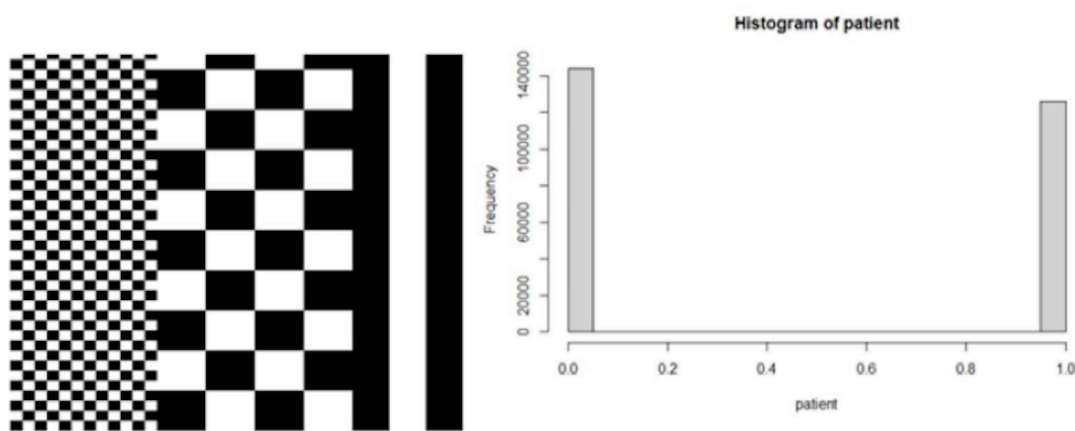


Figura 11. Istogramma dei livelli di grigio. A destra mostra un gran numero di pixel in solo due colonne di intensità di livello di grigio, una per i pixel neri e una per i pixel bianchi.

Derivano dall'istogramma dell'immagine e forniscono metriche come la media, la mediana, i valori massimi e minimi delle intensità dei *voxel*. Comprendono anche caratteristiche come la *skewness* (indicante l'asimmetria), la *kurtosis* (esprimendo l'appiattimento), l'uniformità e l'entropia (indicante la casualità). Ci si aspetta che le caratteristiche di base come massimo, minimo e media corrispondano ai valori di Hounsfield (HU) relativi ai livelli di intensità massima, minima e media, mentre le rimanenti caratteristiche sono essenzialmente misurazioni

statistiche [40].

Le caratteristiche di secondo ordine studiano la *texture* dell'immagine utilizzando i *voxel*. Pertanto, è possibile descrivere le relazioni di similarità tra *voxel* adiacenti.

Le caratteristiche di ordine superiore, d'altra parte, applicano filtri all'immagine, come i frattali, per riconoscere modelli nell'immagine.

2.3. Software utilizzati

La maggior parte delle immagini mediche radiologiche è archiviata in un formato specifico progettato per facilitarne la visualizzazione, la gestione e il trasferimento. Questo formato, ampiamente utilizzato nel settore, è noto come *Digital Imaging and Communication in Medicine* (DICOM).

DICOM è uno standard molto preciso che, ad esempio, assicura la stessa gamma di luminosità su dispositivi diversi mantenendo valori digitali identici, con una risoluzione di 300 DPI e 12 bit di profondità per i livelli di grigio (a differenza degli 8 bit tipici dei dispositivi digitali standard) [41].

Questo formato si basa su una struttura orientata agli oggetti, comprendente un'intestazione e numerosi *tag* che descrivono informazioni come i dati del paziente, i dettagli dell'esame radiologico,

lo scanner e la stampante. Tali elementi collaborano a fornire un quadro radiologico completo. Ogni immagine è rappresentata tramite *tag*, incluso quello relativo ai *pixel* denominato "*image pixel module*". Le immagini DICOM possono essere compresse sia in modalità *lossless* sia *lossy*, e sono visualizzabili esclusivamente tramite *software* compatibili in grado di leggere i metadati. Per ogni esame radiologico si generano numerose immagini DICOM (anche fino a 500), ciascuna delle quali è chiamata "*slice*" [41].

Il formato NRRD (*Nearly Raw Raster Data*) è invece concepito per la rappresentazione e l'archiviazione di dati *raster* multi-dimensionali. Si tratta di un file di testo con due intestazioni che contengono informazioni su spessori, spazi dimensionali e orientamento generale [41].

Con l'avvento della Radiomica, sono state sviluppate migliaia di caratteristiche per analizzare i parametri di Radiomica. Nel 2018, l'*Image Biomarkers Standardization Initiative* (IBSI), frutto di una collaborazione internazionale, ha proposto uno standard per l'estrazione di tali caratteristiche. Il biomarcatore è stato definito come "una proprietà misurata e valutata oggettivamente, utilizzata come indicatore di processi biologici normali, patogeni o risposte a trattamenti farmacologici" [42]. L'IBSI ha inoltre delineato un flusso di lavoro completo per la

Radiomica, fornendo una descrizione formale delle caratteristiche e dei processi di estrazione.

Questo standard è stato integrato nella libreria Python PyRadiomics [43]. PyRadiomics, essendo uno strumento grafico, è incluso nel software open-source 3D Slicer 5.2.2 (Boston, Massachusetts, USA). Quest'ultimo consente non solo la gestione e visualizzazione delle immagini DICOM, ma anche la segmentazione delle ROI, l'estrazione di caratteristiche e altre funzionalità [44].

Il termine "PyRadiomics" deriva dalla combinazione di "Python" e "Radiomics". Si tratta di una libreria progettata per estrarre e analizzare diverse caratteristiche da immagini mediche, come TC e RM, sia in 2D (basate su *pixel*) che in 3D (basate su *voxel*) [43]. Questa libreria è conforme allo standard IBSI, che ha l'obiettivo di promuovere la conoscenza delle funzionalità radiomiche e di ampliare la comunità di utenti.

PyRadiomics offre caratteristiche di primo e secondo ordine e di ordine superiore, organizzate in sette principali categorie (Tabella 8), come segue:

- **FIRST**: acronimo di "Statistiche del primo ordine". Questa categoria descrive la distribuzione delle intensità dei *voxel*

all'interno della ROI utilizzando metriche semplici e comuni, come Media, Mediana ed Entropia.

- **SHAPE**: queste proprietà di ordine superiore non dipendono dai livelli di grigio dell'immagine, ma si basano esclusivamente sulle relazioni spaziali dei *pixel* o *voxel*. Vengono calcolate su una ROI approssimata da una rete triangolare generata tramite l'algoritmo *Marching Cubes* [45].
- **GLCM** (*Gray Level Co-occurrence Matrix*): rappresenta la matrice più utilizzata e diffusa. Misura quante volte si verifica una combinazione di due livelli di grigio i e j in due *pixel* adiacenti dell'immagine.
- **GLDM** (*Gray Level Dependence Matrix*): valuta le dipendenze tra i livelli di grigio nell'immagine, definendo una dipendenza come il numero di *voxel* connessi entro una distanza δ che dipendono dal *voxel* centrale.
- **GLRLM** (*Gray Level Run Length Matrix*): misura le sequenze di livelli di grigio consecutivi, definite come la lunghezza, in *pixel*, di una serie di *pixel* contigui con lo stesso livello di grigio.
- **GLSZM** (*Gray Level Size Zone Matrix*): questa matrice quantifica le zone di livello di grigio in un'immagine. Una zona è costituita da

voxel connessi che condividono la stessa intensità di livello di grigio. La connessione tra *voxel* è determinata da una distanza pari a 1, secondo la regola dell'infinito.

- **NGTDM** (*Neighbouring Gray Tone Difference Matrix*): calcola la differenza tra l'intensità di un livello di grigio e la media dei livelli di grigio dei *voxel* vicini entro una distanza δ . La somma delle differenze assolute per un livello i viene registrata nella matrice.

Per ciascuna di queste matrici, la misura di distanza predefinita è quella euclidea, ma sono disponibili anche altre opzioni, come la distanza di Manhattan o la regola dell'infinito; di default, la GLCM è calcolata in modo simmetrico [46].

FIRST	10 Percentile; 90 Percentile; Energy; Entropy; Interquartile Range; Kurtosis; Maximum; Mean Absolute Deviation; Mean; Median; Minimum; Range; Robust Mean Absolute Deviation; Root Mean Squared; Skewness; Total Energy; Uniformity; Variance.
SHAPE 2D or 3D	Mesh surface; Pixel surface; Perimeter; Perimeter to Surface ratio; Sphericity; Spherical Disproportion; Maximum 2D diameter; Major Axis Length; Minor Axis Length; Elongation.
GLCM	Autocorrelation; Joint Average; Cluster Prominence; Cluster Shade; Cluster Tendency; Contrast; Correlation; Difference Average; Difference Entropy; Difference Variance; Joint Energy; Joint Entropy; Informational Measure of Correlation 1; Informational Measure of Correlation 2; Inverse Difference Moment; Maximal Correlation Coefficient; Inverse Difference Moment Normalized; Inverse Difference; Inverse Difference Normalized; Inverse Variance; Maximum Probability; Sum Average; Sum Entropy; Sum of squares.
GLDM	Small Dependence Emphasis; Large Dependence Emphasis; Gray Level Non-Uniformity; Dependence Non-Uniformity; Dependence Non-Uniformity Normalized; Gray Level Variance; Dependence Variance; Dependence Entropy; Low Gray Level Emphasis; High Gray Level Emphasis; Small Dependence Low Gray Level Emphasis; Small Dependence High Gray Level Emphasis; Large Dependence Low Gray Level Emphasis; Large Dependence High Gray Level Emphasis.
GLSZM	Small Area Emphasis; Large Area Emphasis; Gray Level Non-Uniformity; Gray Level Non-Uniformity Normalized; Size-Zone Non-Uniformity; Size-Zone Non-Uniformity Normalized; Zone Percentage; Gray Level Variance; Zone Variance; Zone Entropy; Low Gray Level Zone Emphasis; High Gray Level Zone Emphasis; Small Area Low Gray Level Emphasis; Small Area High Gray Level Emphasis; Large Area Low Gray Level Emphasis; Large Area High Gray Level Emphasis.
GLRLM	Short Run Emphasis; Long Run Emphasis; Gray Level Non-Uniformity; Gray Level Non-Uniformity Normalized; Run Length Non-Uniformity; Run Length Non-Uniformity Normalized; Run Percentage; Gray Level Variance; Run Variance; Run Entropy; Low Gray Level Run Emphasis; High Gray Level Run Emphasis; Short Run Low Gray Level Emphasis; Short Run High Gray Level Emphasis; Long Run Low Gray Level Emphasis; Long Run High Gray Level Emphasis.
NGTDM	Coarseness; Contrast; Busyness; Complexity; Strength.

Tabella 8. Tabella riepilogativa delle features radiomiche e relative classi usata in Pyradiomics

2.4. Metodi formali

In seguito all'introduzione del Regolamento Generale sulla Protezione dei Dati (GDPR, *General Data Protection Regulation*), che prevede il diritto dell'utente di ottenere una spiegazione chiara e comprensibile della logica alla base dei processi decisionali automatizzati (*explainability*), si è scelto di adottare un approccio alternativo nell'istruzione dei calcolatori, basato sull'impiego dei Metodi formali.

I Metodi formali sono tecniche matematiche e logiche rigorose utilizzate per modellare, analizzare, verificare e validare sistemi di AI o algoritmi di apprendimento automatico. L'obiettivo principale dei Metodi formali è garantire la correttezza e la robustezza di tali sistemi, specialmente in contesti critici dove gli errori possono avere conseguenze gravi, come per l'appunto la diagnostica medica [47] [48].

I Metodi formali consentono l'applicazione di approcci matematici e logici rigorosi per rappresentare e analizzare sistemi complessi, quali modelli predittivi, reti neurali o processi decisionali, in contesti applicativi come la diagnosi medica e l'analisi delle immagini radiologiche. Questi metodi risultano fondamentali per garantire che le decisioni e le analisi condotte dai modelli computazionali siano corrette,

sicure e affidabili. Essi permettono di semplificare le connessioni di un sistema, rappresentandole attraverso una sintassi matematica precisa [47] [48].

Attraverso i Metodi formali, l'apprendimento del sistema computazionale viene realizzato utilizzando una singola immagine, a partire dalla quale si costruisce un modello matematico che consente la verifica di una proprietà definita dall'esperto del dominio, noto come progettista. Nel caso dei Metodi formali applicati alla medicina, l'esperto del dominio è il medico, e la proprietà da analizzare potrebbe riguardare, ad esempio, la presenza o l'assenza di un segno patologico.

Successivamente, il sistema verifica la proprietà e restituisce il risultato: “Vero” se la proprietà è riscontrata nel modello analizzato, oppure “Falso” se il pattern (schema, modello, proprietà o dato) non è presente [47] [48].

Questo approccio garantisce il rispetto GDPR, poiché il modello, oltre a fornire il risultato della verifica, genera un controesempio che illustra la logica impiegata nell'analisi della proprietà. Tale spiegazione supporta la trasparenza e la comprensibilità dei processi decisionali automatizzati.

L'applicazione dei Metodi formali permette di verificare se l'implementazione è allineata alle specifiche e viceversa [47] [48]. La specifica è il processo di descrizione dei requisiti funzionali e non funzionali di un sistema che verrà implementato. La Specifica Formale richiede una sintassi e una semantica definite: una delle più utilizzate è il *Calculus of Communicating Systems* (CCS), proposto da Milner et al. [49].

La rappresentazione grafica del comportamento di un sistema è chiamata Sistema di Transizioni Etichettate (LTS, *Labelled Transition System*) e consiste in stati e transizioni. Il LTS parte da uno stato iniziale, che funge da punto di partenza predefinito; quindi, vengono illustrate le possibili transizioni da uno stato (denominato "P1") a un altro stato (designato come "P2") attraverso un'azione specifica (indicata come "a") [47] [48].

Nel contesto del CCS una serie di operatori e una semantica operativa vengono combinati per costruire automaticamente il modello LTS di un sistema. Questa sintassi operativa ha un grande potere modellistico per la descrizione di un sistema stabilendo assiomi e regole di transizione [47] [48].

Con i Metodi formali diventa possibile “riunire” il campo della scienza informatica con l’esperienza del medico radiologo per costruire sistemi esperti ed efficienti che aiutino nella diagnosi o nella prognosi di varie patologie. Attualmente i Metodi formali rappresentano tra i pochi metodi capaci di risolvere alcune delle limitazioni dell’AI, in particolare la necessità di un volume elevato di dati per garantire la robustezza del modello. Utilizzando i Metodi formali la diagnosi è generata dal model checker, che restituisce un verdetto del tipo “vero” o “falso”, basato sulla soddisfazione o meno di una determinata proprietà della patologia. Quando l’*output* è “falso”, il *Model checker* esprime anche la logica che sta alla base della sua “decisione”, identificando uno specifico punto in cui si verifica la discrepanza tra il modello LTS e la proprietà [47] [48].

L’applicazione dei Metodi formali nella Radiomica si concentra principalmente su tecniche rigorose per la verifica, la validazione e l’interpretabilità dei modelli di machine learning e AI utilizzati per analizzare immagini mediche.

Alcuni esempi di applicazione dei Metodi formali nella Radiomica sono [47]:

- Segmentazione delle immagini: uso di Metodi formali per garantire che gli algoritmi segmentino correttamente tumori e tessuti sani e che piccole variazioni nei dati di *input* (es., rumore) non producano

segmentazioni errate.

- Classificazione tumori benigni/maligni: verifica che le decisioni del modello siano robuste e basate su pattern clinicamente rilevanti, e non su correlazioni spurie nei dati.
- Prognosi basata su biomarcatori di Radiomica: dimostrazione formale del rispetto dei vincoli clinici per l'associazione tra biomarcatori e *outcome* (es., sopravvivenza).

2.5. Tecnica del Model Checking

Il *Model checking* è una tecnica automatizzata per la verifica formale dei modelli di sistemi complessi [50]. Questo metodo permette di verificare che un dato modello di un sistema soddisfi specifiche proprietà espresse nella logica temporale. Tra i contributi fondamentali a questa disciplina ci sono i Metodi formali e il CCS [51]. Questi calcoli permettono di descrivere e analizzare processi concorrenti, fornendo una base solida per la verifica dei sistemi distribuiti. Con i Metodi formali è possibile confrontare comportamenti equivalenti tra modelli; con il *Model checking*, invece, è possibile utilizzare la logica temporale, creando strumenti pratici in contesti reali, come la verifica di hardware e software critici, migliorando la sicurezza e l'affidabilità dei sistemi tecnologici [52] [53].

Gli strumenti del *Model checking* utilizzano la logica temporale per esplorare tutti i possibili stati del sistema e verificare esaustivamente le proprietà. Questi strumenti prendono come *input* sia il Modello formale che le proprietà e producono come *output* un valore binario di giudizio, cioè “Vero” se il sistema soddisfa la proprietà, “Falso” se non la soddisfa [52] [53].

Inoltre, un vantaggio chiave di questi strumenti è la capacità di produrre un controesempio quando il sistema non soddisfa la proprietà, dimostrando così un esempio che falsifica la proprietà verificata. Il controesempio è complesso: può essere utilizzato per identificare e risolvere problemi nel sistema, per comprendere perché il sistema non espone un comportamento specifico, per scoprire ipotesi nascoste e per affinare la comprensione del comportamento del sistema [52] [53].

CAPITOLO III

OBIETTIVO E STRUTTURA DELLO STUDIO

SOMMARIO: 3.1. Endpoint primario dello studio – 3.2 Selezione dei pazienti
– 3.3 Acquisizione delle immagini – 3.4 Segmentazione ed estrazione delle
features – 3.5 Selezione delle features – 3.6 Utilizzo dei metodi formali –
3.7 Risultati – 3.8 Discussione

3.1. Endpoint primario dello studio

L'obiettivo del presente studio è valutare le prestazioni di uno strumento automatizzato per la diagnosi di forme moderate di PA basato sull'analisi delle caratteristiche radiomiche del pancreas. La performance diagnostica dello strumento è stata confrontata con quella di un medico radiologo esperto, utilizzando i criteri CTSI precedentemente descritti come base per la diagnosi radiologica della PA.

Questa ricerca si propone di esplorare le potenzialità della Radiomica attraverso l'applicazione, per la prima volta nel campo della radiologia d'urgenza, dei Metodi formali. In particolare, lo studio mira a ottenere una diagnosi automatizzata di forme moderate di PA mediante l'identificazione di *features* non rilevabili attraverso la sola valutazione radiologica "ad occhio nudo". Tale approccio intende migliorare la performance diagnostica della TC che rimane il *gold standard* nella valutazione della PA.

Sulla base dei risultati ottenuti, è stata inoltre analizzata l'efficacia dello stesso modello nell'anticipare l'evoluzione della PA moderata verso forme gravi, complicate da necrosi. Questo approccio estende il potenziale utilizzo della metodica, non solo in ambito diagnostico, ma anche prognostico, fornendo un ulteriore strumento per la gestione clinica della malattia.

3.2. Selezione dei pazienti

Lo studio è stato condotto in conformità alle linee guida della Dichiarazione di Helsinki e ha ricevuto l'approvazione dall'*Institutional Review Board* dell'Università del Molise (protocollo numero 20/23, data di approvazione: 5 luglio 2023). Sono state raccolte retrospettivamente le scansioni TC con contrasto di 80 pazienti consecutivi, tutti conformi ai criteri di inclusione, eseguite presso la UOC di Radiologia Generale e di PS dell'AORN "Antonio Cardarelli" di Napoli nell'anno 2022 (dal 1 gennaio 2022 al 31 dicembre 2022). Questi pazienti presentavano dolore addominale acuto e un sospetto clinico di PA.

Lo standard di riferimento utilizzato per la diagnosi è stato la PA moderata, definita secondo criteri clinici, laboratoristici e radiologici [16].

I criteri di inclusione sono stati:

- Età superiore ai 18 anni
- TC acquisita secondo protocollo di acquisizione standard predefinito

I criteri di esclusione sono stati:

- Evidenza di lesioni pancreatiche focali o diffuse benigne o maligne
- Diagnosi di PA severa, pancreatite cronica, marcata involuzione della

ghiandola pancreatica

- Mancato consenso a far parte dello studio.

Gli 80 pazienti sono stati divisi in due gruppi.

Il primo gruppo è composto da 40 pazienti che si sono sottoposti alla TC con mdc per dolore addominale acuto ma che non hanno mostrato segni TC di infiammazione a carico della ghiandola pancreatica.

Il secondo gruppo è composto da 40 pazienti con dolore addominale acuto affetti da PA moderata, diagnosticata con TC con mdc in studio multifasico o con criteri clinici e/o di laboratorio.

Gli esami TC sono stati valutati da 3 medici radiologi in cieco e rivalutati da un medico radiologo con 20 anni di esperienza.

3.3. Acquisizione delle immagini

Tutti gli esami TC sono stati eseguiti usando una TC *Dual Source* 256-slice (Somatom Definition Flash; Siemens Healthineers, Erlangen, Germany), con scansioni eseguite in senso cranio-caudale dell'addome in posizione supina, con gli arti superiori addotti. La tecnica *breath-hold* è stata utilizzata per prevenire gli artefatti da movimento e da respiro.

Il protocollo standardizzato TC ha visto l'utilizzo di due fasi ottenute dopo somministrazione di mdc iodato e.v. I parametri TC impiegati sono stati i seguenti: collimazione di 64x0.6 mm, risoluzione spaziale di 0.30 mm, durata di acquisizione di 12 secondi, lunghezza scansione di 287 mm, tempo di rotazione di 0,33 secondi, impostazioni del tubo radiogeno 100/140 kV, 50/60 mAs, dose efficace 2,6 mSv. Queste fasi sono state acquisite dopo intervalli di tempo specifici e hanno anche incluso l'elaborazione di immagini *virtual non-contrast* (VNC) (Tabella 9).

L'acquisizione della fase arteriosa è stata definita temporalmente mediante *bolus tracking* posizionando la ROI a livello dell'arco aortico iniziando l'acquisizione dopo il raggiungimento di una soglia di 150 HU. La fase venosa è stata acquisita con un ritardo di 70-80 secondi dopo l'inizio della somministrazione di mdc iodato e.v. Nella maggior

parte dei casi non si è resa necessaria l'acquisizione di una fase tardiva aggiuntiva a 3-5 minuti, salvo che in casi selezionati. Le immagini VNC sono state ottenute in *post processing* sottraendo il valore di HU derivante dallo iodio al valore di HU totale di ogni singolo *voxel*, permettendo di ottenere così una ricostruzione virtuale dell'immagine "pre-contrastografica" e una globale riduzione della dose di esposizione al paziente. Il volume acquisito ha compreso la regione addominale dal piano diaframmatico alle branche pubiche, con uno spessore di fetta di 2.5 mm e uno spessore di ricostruzione di 0.625 mm.

Parametri di acquisizione TC	
Collimazione	64 x 0.6mm
Risoluzione spaziale	0.30 mm
Scan time	12 s
Scan length	287 mm
Tempo di rotazione	0.33 s
Tube settings	100/140kV , 50/60 mAs
Dose efficace	2.6 mSV
Protocollo di somministrazione del mezzo di contrasto	
Alta concentrazione di mdc e.v. (Iomeprol 400mg I/ml, Iomeron 400; Bracco, Milano, Italia)	1.2 ml per kg di peso corporeo
Flusso	3.5 - 4 mL/s
ROI	150 HU
Ritardo fase arteriosa	15 s
Ritardo fase portale	70-80 s

Tabella 9. Protocollo di acquisizione TC

3.4. Segmentazione ed estrazione delle features

Le immagini analizzate nel presente studio sono state acquisite durante la fase venosa, con uno spessore di fetta pari a 2,5 mm. La scelta di questa specifica fase di acquisizione è motivata dalla sua capacità di ottimizzare la risoluzione del contrasto tra il tessuto ghiandolare sano e la reazione infiammatoria intra- e peri-ghiandolare.

La ROI è stata definita a mano su ogni esame da 6 diversi operatori, fetta per fetta, per ogni paziente, utilizzando il software open source 3D slicer versione 5.2.2 (Boston, Massachusetts, USA). Prima di cominciare la segmentazione manuale, è stato definito un protocollo per evitare la variabilità inter-operatore. La ROI ha incluso esclusivamente la ghiandola pancreatica, evitando rigorosamente di includere i vasi principali (vena porta, vena mesenterica, vena splenica), il duodeno e il piccolo intestino. Quando presente del fluido peri-pancreatico, quest'ultimo è stato invece incluso nella ROI (Figura 12).

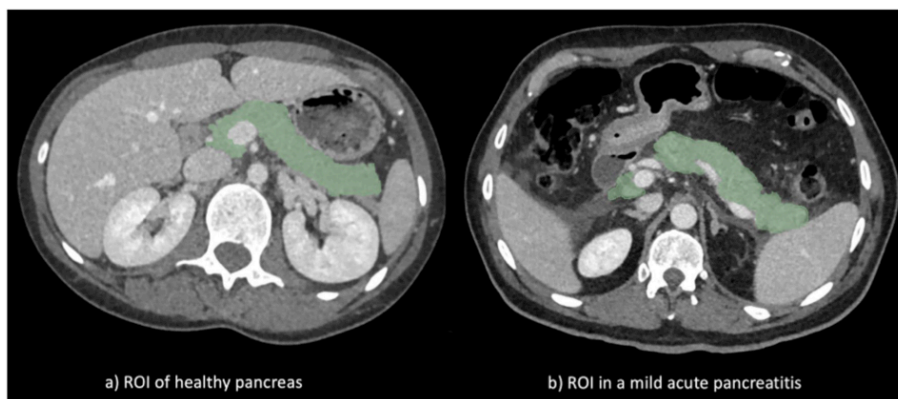


Figura 12. a) ROI di un paziente con pancreas sano; b) ROI in un paziente con PA. ROI (tracciata in verde)

In media, il numero di ROI estrapolate nei pazienti sani è stato pari a 19,7 (intervallo 6-27, intervallo di confidenza al 95%), mentre nei pazienti affetti da patologia il numero medio di ROI estrapolate è risultato pari a 22,7 (intervallo 19-31, intervallo di confidenza al 95%).

Successivamente, le *features* radiomiche sono state estratte utilizzando il software “PyRadiomics” (Boston, Massachusetts, USA). Come precedentemente descritto, *PyRadiomics* è un’estensione del software 3D Slicer versione 5.2.2, nonché una libreria Python che consente di convertire le immagini radiologiche in formato DICOM in dati numerici. Questo processo permette di trasformare le immagini radiologiche in *features*, rappresentandole come valori numerici [54].

3.5. Selezione delle features

Le 107 features estratte attraverso PyRadiomics sono state visualizzate e poi selezionate tramite Weka (General Public License, New Zealand), un software open-source che, attraverso una serie di algoritmi di machine learning e “*data mining*”, è capace di identificare le features più discriminanti in un dataset in formato CSV (comma-separated value) (Figura 13).

firstorder_10Percentile	firstorder_Energy	firstorder_Entropy	firstorder_Kurtosis	firstorder_Maximum	firstorder_Mean	firstorder_Median	firstorder_Minimum	firstorder_Range	firstorder_Skewness	firstorder_TotalEnergy	firstorder_Uniformity	firstorder_Variance	class
58.43684211	115852016.1	1.815144875	4.631221974	161.4605263	85.29174078	87.17761158	-30.72368421	192.1842105	-0.62198074	192980156.4	0.347824478	478.2028628	malati
112.2666667	310123912	1.731514998	6.734632923	245.3015873	136.6636684	137.8492063	8.936507937	236.3650794	-0.005601133	347502418	0.387661004	451.3351512	malati
97.1625	211561199.3	1.689959149	5.291615219	209.5	120.4085934	121.35	0.6175	208.8625	-0.382001003	298963418.1	0.377727425	395.828912	malati
73.92	139543424.7	1.695620922	6.086018292	168.86	99.27281818	101.64	19.22	186.08	-0.940253853	165689307.4	0.401037881	466.8763159	malati
61.6888889	217225484.9	1.912468152	184.7188725	918.1746012	94.56013076	97.49206349	-73.82539683	992	6.24760503	264678925.1	0.33691974	1136.684976	malati
120.5873016	255643931	1.672076305	6.181889325	243.4126984	143.7364052	144.6111111	35.9047619	207.5079365	-0.162075112	317912279.8	0.385960727	393.8454726	malati
88.3115942	130962871.6	1.900602543	5.533974446	227.6211884	115.4389832	115.4710145	-5.869565217	233.4927536	0.145175332	188054144	0.332859721	534.8746145	malati
105.9555556	161557350.6	1.89120005	3.952632044	227.1313333	133.4088891	134.2777778	21.55555556	205.5777778	-0.318605588	374397623	0.321691112	472.9130906	malati
79.8179773	101613390	1.595813818	5.042177297	180.5955056	100.9391954	101.4048444	1.921348115	178.6745373	-0.200744162	189151903.4	0.39792359	310.012351	malati
136.2240964	281767190.6	1.39814584	5.222036055	233.6024096	154.5748285	154.9759036	70.6626506	162.939759	-0.056194361	304023817.8	0.474202709	242.087993	sani
97.1079494	130672416.3	1.51943664	14.2820692	203.2278481	116.5952889	116.8860759	0.46835443	202.7594937	-0.476223901	219580743.2	0.417333305	367.0836204	sani
55.04545455	120450813.3	1.638581649	3.912566526	159.3181818	77.36568506	77.93939394	-11.93939394	171.2575758	-0.143820787	134247675.1	0.386858901	321.3040107	sani
73.94805195	108768321.8	1.445717079	4.968131949	152.3116883	93.02458033	94.02397403	1.116883117	151.1948052	-0.48991895	189214180	0.42882738	258.5995083	sani
81.62444444	21490613.2	1.554061692	5.378575135	207.9	103.9601712	104	12.64444444	195.2555556	0.241332863	181794736.5	0.418176143	292.1526155	sani
76.48795181	158433436.2	1.58433284	4.430615581	175.1927711	96.75286052	99.8493979	5.228915663	169.9638554	-0.238516139	203070200.2	0.401131736	324.2441824	sani
68.78873239	97564544.07	1.446270147	4.189842007	153.4788732	87.57767666	87.90140845	4.661971831	148.8169014	-0.178972457	132667890.7	0.445550839	234.9718611	sani
88.2797619	165802428.7	1.621762457	5.16302312	200.9642857	110.6917139	111.1428071	2.464285714	198.5	-0.093891936	261840125.8	0.400380644	362.2792902	sani
99.33333333	145168119.9	1.485714498	5.723002738	202.1666667	118.4509323	118.775641	29.93589744	172.2307692	0.189089945	243939192	0.428996054	286.3173441	sani
67.38235294	155731545.3	1.801874502	4.079327316	182.5098039	93.51779995	94.37254902	-15.49019608	196	-0.195485735	236869596.2	0.349927424	440.7165541	sani
12.26060606	14162823.08	1.494773849	8.000827767	146.5353535	30.51642801	28.75757576	-41.03030303	187.5656566	1.181684054	24743604.42	0.424860835	534.9514552	sani
73.30757576	151566471.5	1.676371099	4.078513052	191.9848485	96.18747429	96.34848485	13.43939394	178.5454545	0.070597065	232932846.5	0.371811164	339.4487461	sani
106.6587568	248740468.4	1.505303805	5.564919067	208.5540541	126.4850885	126.7364865	33.36486486	175.1891892	-0.1527414	315639949.1	0.431567257	271.5001549	sani
64.24166667	879773714.61	1.816449311	5.421136283	181.5416667	91.23082835	92.54861111	-35.66666667	217.2083333	-0.336290235	172802082.7	0.34832173	512.8546173	sani
82.3182958	129656219.9	1.583699602	5.353777357	178.0123951	101.0541855	103.5731707	-5.975609756	183.5878049	-0.449213688	239586869.2	0.407584893	302.992465	sani
88.86390476	141248433.1	1.497514683	9.139444134	176.1666667	108.909921	110.5	-38.63095238	208.797619	-1.230382616	194917762.3	0.441021893	314.3959609	sani
58.92525253	93782190.88	1.811020286	5.182335489	181.0310101	84.56686967	85.18161616	-11.60060601	192.6161616	-0.102731302	187779065.9	0.342739887	462.1726077	sani
97.52985075	170968066.1	1.485649931	15.9460404	207.9402985	117.6928258	118.9850746	26.13432836	181.8059701	0.02037905	212611531.4	0.427075857	300.8211047	sani
102.2146341	206460358.3	1.601342252	9.251837459	233.804878	123.051404	123.1097561	-12.46341463	246.2682927	0.158827979	228876882.5	0.410602325	374.4154521	sani
59.05128205	119380102.3	1.597844138	6.767021985	177.3205128	81.23471553	82.43589744	-40.93589744	218.2564103	-0.59818616	181765532.7	0.415803385	366.7252044	sani
45.62325581	70380980.02	1.769072943	6.099699523	181.8139535	69.17428991	66.37293932	-33.34883723	215.1627907	0.518243599	104951337.8	0.365385337	433.1946488	sani

Figura 13. Esempio di file formato CSV

Nel nostro studio, il software Weka ha selezionato le features più discriminanti e significative relative ai pancreas “sani” e “malati”, mediante una *feature selection* tramite gli algoritmi “*Gain Ratio Attribute Evaluation*” e “*Correlation Attribute Evaluation*” dipendentemente dai risultati di una successiva classificazione supervisionata, ritenuta appropriata per valori di *Precision* e *Recall* superiori all’80% dove possibile; tale selezione ha permesso di ottenere 22 *features* suddivise in 6 classi [55] (Tabella 10):

- *First order features*: le *features* di questa classe analizzano la distribuzione e la frequenza dei valori assunti dai *voxel* presenti nella ROI.
 - *10th percentile*, indice del 10° percentile di X (dove X indica un set di N *voxel* inclusi nella ROI;

- *Mean*, indice del valore medio di intensità di grigio entro la ROI;

$$mean = \frac{1}{N_p} \sum_{i=1}^{N_p} \mathbf{X}(i)$$

- *Median*, indice del valore mediano di intensità di grigio entro la ROI;

- *Skewness*, che misura l'asimmetria della distribuzione dei valori nell'ambito del valore medio (Mean value), può essere un valore positivo o negativo;

$$skewness = \frac{\mu_3}{\sigma^3} = \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^3}{\left(\sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^2} \right)^3}$$

μ_3 rappresenta il 3° momento centrale

- *Firstorder Uniformity*, indice dell'uniformità dei valori dell'istogramma;

$$uniformity = \sum_{i=1}^{N_g} p(i)^2$$

- *Grey Level Dependence Matrix Features*: le features di questa classe analizzano la distribuzione dei valori della matrice GLDM, ottenuta ponendo come valore dell' (i, j) - esimo elemento della matrice, la frequenza con cui un *voxel* di valore i è circondato da *voxel* di valore j dipendente da i ; la dipendenza è verificata se $|i - j| < \alpha$ (spesso considerato uguale a 0).
 - *Gldm Dependence Entropy*, indice della non uniformità della matrice;

$$DependenceEntropy = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i, j) \log_2(p(i, j) + \epsilon)$$

- *Gldm Low Gray Level Emphasis*, indice della distribuzione di valori bassi di grigio nell'immagine;

$$LGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{P(i, j)}{i^2}}{N_z}$$

- *Gldm Dependence non uniformity*, indice della somiglianza di dipendenza nel contesto dell'immagine, con un valore basso che sta ad indicare maggiore omogeneità tra le dipendenze nell'immagine;

$$DN = \frac{\sum_{j=1}^{N_d} \left(\sum_{i=1}^{N_g} P(i, j) \right)^2}{N_z}$$

- *Grey Level Co-occurrence Matrix Features (GLCM simmetrica)*: le *features* di questa classe analizzano la distribuzione dei valori della matrice GLCM, ottenuta ponendo come valore dell' (i, j) -esimo elemento della matrice, la frequenza con cui un *voxel* di valore i è affiancato da un *pixel* di valore j .

- *Gldm Imc2 (Informational Measure of Correlation)*, indica la correlazione tra le distribuzioni di probabilità di i e j , quantificando la complessità della texture;

$$IMC\ 2 = \sqrt{1 - e^{-2(H_{XY2} - H_{XY})}}$$

- *GlcM Maximum Probability*, rappresenta l'occorrenza della coppia predominante di valori di intensità vicini;

$$\text{maximum probability} = \max (p(i, j))$$

- *GlcM Joint Entropy*, indice della casualità dei valori di grigio vicini;

$$\text{joint entropy} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log_2 (p(i, j) + \epsilon)$$

- *GlcM Sum Entropy*, esprime la somma di *voxel* vicini con intensità differente;

$$\text{sum entropy} = \sum_{k=2}^{2N_g} p_{x+y}(k) \log_2 (p_{x+y}(k) + \epsilon)$$

- *Grey Level Run Length Matrix Features*: le *features* di questa classe analizzano la distribuzione dei valori della matrice GLRLM, ottenuta ponendo come valore dell' (i, j) -esimo elemento della matrice, la frequenza con cui un *voxel* di valore i si ripete per j volte.

- *Glrlm Low Gray Level Run Emphasis*, indice della distribuzione di bassi livelli di grigio;

$$LGLRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j|\theta)}{i^2}}{N_r(\theta)}$$

- *Glrlm Short Run Low Gray Level Emphasis*, indice della distribuzione di corte sequenze con bassi livelli di grigio;

$$SRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j|\theta)}{i^2 j^2}}{N_r(\theta)}$$

- *Glrlm Gray Level non uniformity*, indice della variabilità dei valori di intensità di grigio nell'immagine, con un valore tanto più basso quanto maggiore è l'omogeneità tra i valori;

$$GLN = \frac{\sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_s} P(i, j) \right)^2}{N_z}$$

- *Glrlm Run entropy*, che indica la casualità/incertezza nella distribuzione delle “*run lengths*” e dei livelli di grigio, con un valore tanto più alto quanto maggiore è l’eterogeneità nei pattern della *texture*;

$$LGLRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i, j | \theta)}{i^2}}{N_r(\theta)}$$

- *Grey Level Size Zone Matrix Features*: le *features* di questa classe analizzano la distribuzione dei valori della matrice GLSZM, ottenuta ponendo come valore dell’ (i, j) -esimo elemento della matrice, la frequenza con cui una zona di valore i e grandezza j è presente nell’immagine. La grandezza è definita considerando la normale uniforme al *voxel*, sono quindi considerati vicini ad ogni elemento 8 *voxel* in 2D e 26 *voxel* in 3D.

- *Glszm Small Area High Gray Level Emphasis*, indice della distribuzione di zone più piccole con livelli più elevati di grigio;

$$SALGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i,j)}{i^2 j^2}}{N_z}$$

- *Glszm Size Zone Non Uniformity*, indice della variabilità dei “size zone volumes” nell’immagine, con un valore tanto più basso quanto maggiore è l’omogeneità;

$$SZN = \frac{\sum_{j=1}^{N_s} \left(\sum_{i=1}^{N_g} P(i,j) \right)^2}{N_z}$$

- *Glszm Zone Entropy*, che indica la casualità /incertezza nella distribuzione delle “zone sizes” e dei livelli di grigio, con un valore tanto più alto quanto maggiore è l’eterogeneità nei pattern della *texture*;

$$ZE = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j) \log_2(p(i,j) + \epsilon)$$

- *Shape features*: le *features* di questa classe analizzano i tre diametri massimi e la forma della ROI, derivate dalla forma approssimativa definita dalla “mesh” derivata della ROI. Esse sono quindi indipendenti dai livelli di grigio.
 - *Shape Maximum 2D Diameter (Row)*, indice della distanza massima euclidea tra coppie di vertici della mesh della lesione all’interno del piano (solitamente sagittale);
 - *Shape Minor Axis Length*, indice della lunghezza del secondo asse più grande dell’ellissoide che racchiude la ROI;

$$minor\ axis = 4\sqrt{\lambda_{minor}}$$

- *Shape Surface Area*, calcolata sommando le aree dei singoli triangoli che compongono la mesh;

$$A_i = \frac{1}{2} |\mathbf{a}_i \mathbf{b}_i \times \mathbf{a}_i \mathbf{c}_i| \quad (1)$$

$$A = \sum_{i=1}^{N_f} A_i \quad (2)$$

FIRST	GLCM	GLDM
10Percentile	Imc2	Dependence entropy
Mean	Joint entropy	Low gray level emphasis
Median	Sum entropy	Dependence non uniformity
Skewness	Maximum probability	
Uniformity		
GLSZM	SHAPE	GLRLM
Zone Entropy	Maximum 2D Diameter Row	Gray level non uniformity
Small area high gray level emphasis	Minor axis length	Low gray level run emphasis
Size zone non uniformity	Surface area	Run entropy
		Short run low gray level emphasis

Tabella 10. In elenco le 22 features estratte divise in 6 classi

3.6. Utilizzo dei metodi formali

Una volta scelte e rese disponibili le *features* più adatte e significative per il nostro caso di studio, sono stati creati dei modelli appropriati utilizzando Metodi formali discretizzando i valori ottenuti (ovvero convertendo l'insieme di valori continui ottenuti in un insieme discreto) e traducendoli nella sintassi matematica CCS.

Tramite i Metodi formali si ha la possibilità di rappresentare ogni singolo paziente con un LTS ovvero un sistema di transizione etichettato, quindi con un sistema a transizione di stati che possono essere rappresentati, partendo da uno stato iniziale, attraverso degli archi orientati ed etichettati.

Sono stati ottenuti cinque intervalli, definiti rispettivamente “*very low*”, “*low*”, “*basal*”, “*high*” e “*very high*”; la presenza o assenza di

suddetti intervalli nella serie di immagini riferita ad un paziente è fondamentale per la classificazione. In questo tipo di sintassi matematica non c'è necessità di un *testing set* o di un *training set*, ma al momento della classificazione c'è bisogno di una “*knowledge extraction*”, ovvero di una estrazione di conoscenza da un esperto del dominio, nel nostro caso il medico radiologo.

Per differenziare un pancreas sano da una PA, due esperti medici radiologi hanno selezionato 6 pazienti affetti da PA moderata con uno score di CTSI di 3 per creare una proprietà formale (ovvero la “regola”) per poter eseguire il lavoro di classificazione. Questi pazienti sono stati selezionati dal nostro database tenendo conto dell'evidenza dell'edema pancreatico, della densità della ghiandola o dell'eventuale involuzione ghiandolare o presenza di fluido peri-pancreatico.

Le regole vengono definite per discriminare la presenza di features radiologiche di PA moderata, così che non ci sia il bisogno di definire allo stesso modo regole per il caso dei pancreas sani, che in tal modo fungono esclusivamente da gruppo di controllo per la validazione.

La proprietà (ovvero la “regola”) è stata verificata tramite *Model checking*, con l'utilizzo di una tecnica formale matematica utile a valutare e controllare se una definita proprietà è soddisfatta o meno da ognuno dei pazienti facenti parte della residua coorte [56] [57] [58].

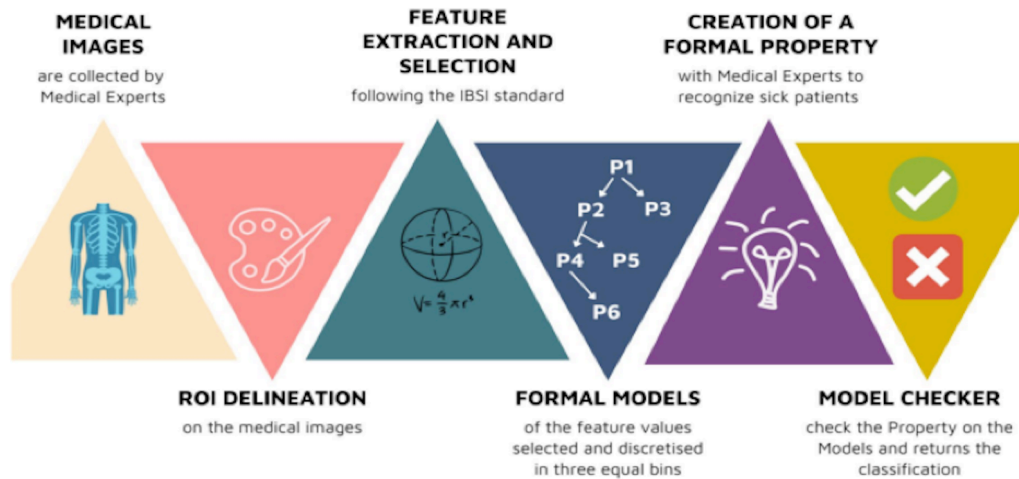


Figura 14. Schema della metodologia impiegata: a partire dalle immagini mediche c'è il passaggio alla segmentazione delle ROI definendo e selezionando la regione pancreatica. Dopo l'estrazione e la selezione delle features, vengono creati modelli matematici e successivamente verificati attraverso la proprietà formale definita tramite la collaborazione tra medici radiologi ed informatici. Il Model checker controlla successivamente se la proprietà è soddisfatta o no, definendo il paziente rispettivamente come sano o affetto da PA

3.7. Risultati

Durante l'analisi dei dati, le *features* estratte da 4 pazienti (di cui 1 sano e 3 con PA moderata) sono risultate incomplete; per cui tali pazienti sono stati esclusi dal gruppo di studio. In tutto sono stati considerati quindi 76 pazienti, di cui 39 soggetti sani (controlli) e 37 pazienti con PA moderata (casi). Quest'ultimo gruppo include i 6 pazienti con CTSI di 3 selezionati per definire la “*Formal Property rule*”.

La regola più performante è stata ottenuta da 3 dei 6 pazienti, precedentemente valutati dai medici radiologi, ed è stata definita a partire dalle classi FIRST e SHAPE selezionate dai seguenti pazienti: FIRS CSV_20_ CSV35 OR SHAP_CSV20_CSV29.

La regola è stata successivamente estesa ed applicata ai restanti 70 pazienti. I risultati sono riportati nella Tabella 11.

Matrice di confusione	Valori effettivi	
Valori previsti	TP 25	FP 7
	FN 6	TN 32
Performance diagnostica n=70	Accuracy	0.81
	Precision	0.78
	Sensitivity	0.81
	Specificity	0.84

Tabella 11. Risultati del Model checker nel dataset in esame (n=70); TP veri positivi, FP falsi positivi, FN falsi negativi, TN veri negativi

32 pazienti sani sono stati correttamente classificati come negativi tramite i Metodi formali.

L'81% dei pazienti con PA moderata è stato inquadrato autonomamente dal modello di Radiomica.

Altri medici radiologi in cieco hanno valutato gli esami ed è stata verificata la loro concordanza inter-operatore. 6 pazienti con PA moderata e CTSI di 0-1 sono stati erroneamente diagnosticati come sani dai medici radiologi (Figura 15a), solo 2 di questi 6 pazienti sono stati erroneamente diagnosticati anche dal modello di Radiomica. 2 giovani pazienti sono stati “sovra-diagnosticati” come PA moderata dai medici radiologi e correttamente classificati invece grazie ai Metodi formali (Figura 15b).

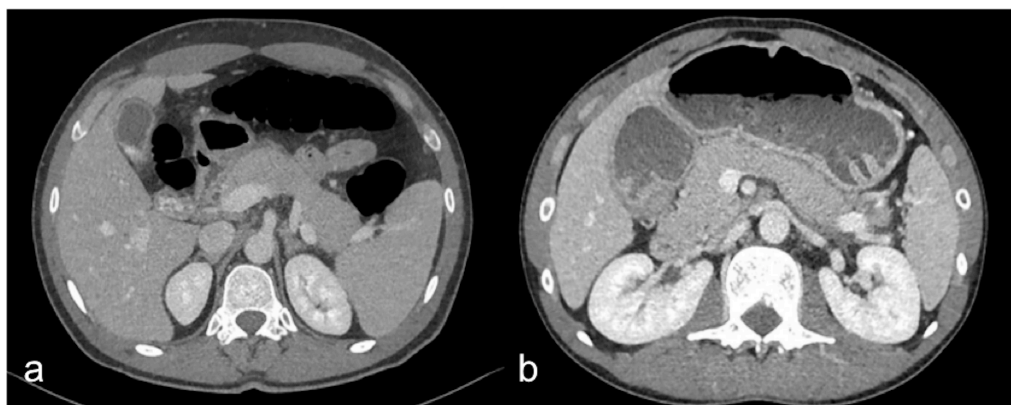


Figura 15. a. Esempio di un giovane paziente affetto da PA moderata, erroneamente classificato come sano dal medico radiologo, ma correttamente identificato come positivo grazie all'impiego dei Metodi formali.
b. Esempio di un giovane paziente sano, diagnosticato erroneamente come PA moderata dal medico radiologo, ma correttamente classificato come sano tramite l'utilizzo dei Metodi formali.

La performance diagnostica raggiunta dal modello di combinazione di Radiomica e valutazione radiologica è riportata in Tabella 12

Matrice di confusione	Valori effettivi	
Valori previsti	TP 29	FP 0
	FN 2	TN 39
Performance diagnostica n=70	Accuracy	0.97
	Precision	1.00
	Sensitivity	0.94
	Specificity	0.95

Tabella 12. Performance diagnostica del modello combinato di valutazione radiologica e Metodi formali

Le ROI sono state selezionate da 6 differenti operatori; questo processo non ha condizionato l'estrazione delle features, la selezione delle features né la verifica formale.

3.8. *Discussione*

Il nostro studio ha dimostrato come i Metodi formali siano affidabili nel riconoscere in breve tempo una PA moderata senza la necessità di dati clinici né dell'intervento "umano" del professionista radiologo dopo opportuna definizione della metodologia.

In particolare, si è visto come siano stati in grado di ottenere un'elevata sensibilità nel riconoscere la PA moderata in pazienti con CTSI di 0 o 1. Il modello ha raggiunto una grande accuratezza nel differenziare entrambi i gruppi di pazienti. In particolare:

Il modello ha raggiunto un'elevata accuratezza nella differenziazione tra i due gruppi di pazienti. In particolare, i pazienti sani erroneamente classificati anche dai medici radiologi come affetti da PA erano giovani (età < 40 anni), normopeso o sovrappeso, con una ghiandola pancreatica ipertrofica situata in uno spazio retroperitoneale ridotto e una bassa percentuale di grasso circostante. Inoltre, anche i pazienti con PA classificati erroneamente presentavano una densità ghiandolare diversa, attribuibile a componente fibrotica. Infatti, gli unici pazienti che presentavano queste caratteristiche differenti sono stati quelli considerati dal modello come falsi negativi, ma correttamente classificati dai medici radiologi.

Secondo la nostra opinione l'attendibilità del modello potrebbe dare

ai medici radiologi e clinici la possibilità di usufruire di questa ulteriore arma durante la pratica clinica. L'accuratezza globale, infatti, risulta comparabile con quella del medico radiologo, ma l'accuratezza globale del modello in combinazione con l'*expertise* del medico radiologo potrebbe permettere di raggiungere un'accuratezza di quasi il 100% , nel riconoscere tanto i pazienti con CTSI 0-1 quanto quelli con CTSI 2-3.

Inoltre, nella pratica clinica una diagnosi precoce deve rappresentare l'obiettivo principale per i medici clinici e radiologi, poiché permette una caratterizzazione accurata e mirata della PA, fondamentale per poter migliorare la prognosi del paziente e ridurre quindi la mortalità.

Il CTSI, che ricordiamo essere basato sui reperti delle immagini TC, può identificare il corretto stadio della pancreatite soltanto a partire dalle 72 ore dall'insorgenza dei sintomi, in quanto vi è un intervallo temporale riconosciuto tra l'insorgenza dei sintomi, dunque del fenomeno patologico, e la comparsa di segni radiologici [15] [16] [17].

In tale contesto, il nostro modello potrebbe invece identificare i più precoci segni TC della PA moderata secondo un nuovo score di severità “radiomico” e non più puramente radiologico [59].

Secondo la nostra esperienza, questo lavoro rappresenta il primo studio basato su un algoritmo creato a partire dalle *features* estratte da 6 pazienti con imaging “tipico” di PA, selezionati e scelti da due medici

radiologi di comune accordo durante la valutazione delle features radiologiche.

L'innovazione del metodo ibrido applicato ci ha permesso di ottenere una elevata attendibilità giacché la regola diagnostica non si è basata solo sulle *features* valutate con l'osservazione “a occhio nudo”, ma anche sulla conversione delle features radiologiche in valori numerici [60] [61] [62].

Un'altra innovazione dell'utilizzo dei Metodi formali, nonostante l'utilizzo delle più comuni metodiche di *Machine Learning*, consiste nell'aver superato i limiti della necessità di un ampio dataset e di un accurato sistema di addestramento [63] [64].

Come descritto, i modelli di *Machine Learning* sono strutturati mediante un set di addestramento, un set di validazione e un set di “*testing*”, dato che il software apprende dati dalle immagini radiologiche. Le immagini possono inoltre essere associate a dei dati clinici che possono essere inseriti nel contesto del processo di apprendimento del modello.

Per questo motivo, il modello applicato è stato sviluppato senza il ricorso a modelli di *Machine Learning*. La regola è stata manualmente definita su 3 dei 6 pazienti affetti da PA moderata, considerando le immagini più “tipiche” della patologia in esame da due medici radiologi esperti.

Tralasciando il passaggio precedentemente esaminato relativo al set di addestramento, i restanti pazienti selezionati sono stati impiegati esclusivamente per la verifica della regola (training/testing 6/70), consentendo al sistema di classificare correttamente l'81% dei casi.

Inoltre, l'uso dei Metodi formali si è mostrato affidabile nell'analizzare un dataset di immagini nonostante la selezione delle ROI da parte di 6 diversi operatori. Questo processo infatti non ha condizionato l'estrazione delle *features* e i risultati, per cui la segmentazione può considerarsi in tal caso operatore-indipendente.

Quindi, oltre alle altre potenziali applicazioni dei Metodi formali nella pratica clinica, l'applicazione alla PA mostra diverse opportunità per l'applicazione dell'AI rispetto a quanto non sia stato fatto negli studi condotti finora. Secondo la nostra opinione, la possibilità di classificare i pazienti le cui segmentazioni vengono eseguite manualmente da diversi operatori rappresenta uno dei principali e più significativi obiettivi del nostro studio.

Inoltre, secondo quanto ci risulta, si tratta anche del primo studio riguardante la diagnosi automatizzata di PA mediante l'analisi della sola fase portale, il che rappresenta un forte punto a favore dato che riduce il tempo necessario alla segmentazione conservato il miglior *enhancement* possibile della ghiandola pancreatica [62] [65] [66].

Si tratta inoltre di uno studio innovativo che mette a confronto

diverse applicazioni della Radiomica alle più recenti innovazioni nel campo tecnologico della TC.

Gli studi di Radiomica applicati alla medicina e alla radiologia d'urgenza si sono finora focalizzati sulla previsione della prognosi, sull'identificazione e sulla diagnosi della pancreatite cronica, oltre che sulla diagnosi differenziale tra adenocarcinoma pancreatico e pancreatite autoimmune [62] [67].

Sebbene diversi studi riportino metodi applicati anche ad un elevato numero di pazienti, spesso questi non hanno utilizzato un gruppo di pazienti sani di controllo, concentrandosi piuttosto soltanto sul gruppo di pazienti malati affetti da PA per stimarne la prognosi [68] [69] [70].

La natura retrospettiva del nostro studio ne rappresenta il principale limite. Il numero di pazienti è ancora limitato, ma comunque si è mostrato sufficiente nell'impostare la regola che è stata “addestrata” e validata.

Inoltre abbiamo escluso dal nostro dataset casi di pancreatite cronica, sebbene possa essere una causa di dolore addominale acuto in caso di riattivazione.

Questo studio rappresenta un primo passo non solo per verificare l'attendibilità del metodo ibrido applicato creato da medici radiologi, clinici ed informatici, ma anche per rendere semi-automatizzata la diagnosi di PA.

Il principale obiettivo si rivolge alla gestione ospedaliera di pazienti ad alto rischio di PA, per evitare l'insorgenza di complicanze, in particolare in caso di evoluzione in PA necrotizzante.

La PA necrotizzante infetta rappresenta un quadro clinico molto complesso da gestire e solitamente può richiedere l'impiego di un approccio chirurgico o interventistico da parte di personale qualificato ed esperto.

Nonostante il crescente utilizzo di tecniche interventistiche sempre meno invasive, la letteratura attuale tende a favorire il drenaggio e la necrosectomia endoscopica, ove applicabili [71]. Tuttavia, la gestione ottimale della PA necrotizzante resta oggetto di dibattito e dovrebbe essere personalizzata in base alle condizioni cliniche del paziente, alle comorbidità e al livello di specializzazione del centro sanitario.

CAPITOLO IV

PROSPETTIVE FUTURE

Sulla base dei promettenti risultati dello studio oggetto di questa Tesi, in cui un modello sviluppato mediante l'utilizzo dei Metodi formali ha dimostrato un'elevata efficacia nel riconoscere casi di PA moderata, raggiungendo circa l'80% di sensibilità, specificità e accuratezza, è stato condotto un ulteriore approfondimento. Questo aveva l'obiettivo di perfezionare il modello, consentendo l'identificazione di pazienti a più alto rischio di progressione della malattia e con una prognosi più sfavorevole.

L'identificazione precoce della progressione della PA moderata verso forme severe o necrotizzanti riveste un ruolo cruciale, in quanto consente di attuare interventi clinici tempestivi, ridurre la mortalità e ottimizzare l'impiego delle risorse nella gestione di questi pazienti [72] [73].

Nello studio è stata inclusa una coorte di 41 pazienti con diagnosi di PA moderata all'ingresso, di cui 37 rappresentano il sottogruppo dei casi analizzati nel precedente studio, mentre ulteriori 4 pazienti sono stati selezionati estendendo retrospettivamente la raccolta dei dati anche al mese di Gennaio 2023 (dal 1 gennaio 2023 al 31 gennaio 2023) seguendo la stessa metodologia e gli stessi criteri di inclusione dei precedenti, con l'obiettivo di ampliare il campione di pazienti con PA moderata all'ingresso. I dati radiologici sono stati analizzati mediante l'utilizzo del *Model checking*, integrato con un *Reduction Step* per

gestire il fenomeno della “*State explosion*” (esplosione degli stati).

Questa strategia di *Model building* può risultare un utile strumento per i medici, da un lato, fornendo un aiuto nel costruire piani di gestione sempre più personalizzati sulla base del singolo paziente, e per gli informatici del campo medico, migliorando la conoscenza di questo tipo di patologia e dell'applicazione della Radiomica nella sua gestione all'interno della comunità scientifica [72] [73].

Il metodo utilizzato (Figura 16) è il medesimo fin qui discusso, nonché il processo essenziale della metodologia applicata basata sul *Model checking*. Come già precedentemente descritto, la tecnica del *Model checking* richiede due *input* distinti: un Modello formale del sistema e un insieme di Proprietà che il sistema deve verificare. Più precisamente, il Modello formale imita il comportamento del sistema ed è generalmente rappresentato come un automa, che include un insieme di nodi insieme a un insieme di archi etichettati che collegano questi nodi tra loro. Un nodo denota uno stato del sistema e un arco etichettato rappresenta una transizione tra gli stati del sistema. Ad esempio, dato l'arco $s \rightarrow s'$, il sistema può passare dallo stato s allo stato s' eseguendo l'azione a .

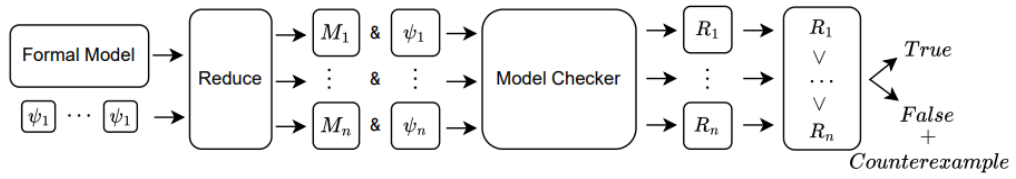


Figura 16. Workflow generale della metodologia basata sul Model checking

I Metodi formali affrontano il problema dell'esplosione degli stati, dove il numero di stati e transizioni diventa troppo grande per essere modellato quando si trattano sistemi complessi, ad esempio, immagini mediche [74]. Numerose tecniche sono state proposte nella letteratura per affrontare questo problema. Anche questo lavoro ha visto l'applicazione di una tecnica di riduzione basata su formule logiche; più in dettaglio, è stato utilizzato il *mu-calculus* selettivo [75] [76], ovvero una logica temporale che consente di ignorare le azioni del sistema in base alla proprietà da verificare. Si noti che, attraverso il *mu-calculus* selettivo, le azioni rilevanti per la verifica di una formula sono solo quelle esplicitamente specificate negli operatori modali della formula.

Pertanto, per ridurre la dimensione dello spazio degli stati, è possibile minimizzare l'automa eliminando tutte le azioni non rilevanti per la proprietà specificata.

Nello studio si è integrata la metodologia basata sul *Model checking* nell'analisi delle immagini mediche. In particolare, le procedure che hanno reso possibile questa integrazione nell'ambito

dell'analisi delle immagini mediche sono state le seguenti:

- Composizione parallela

Essa prende come *input* tutte le tracce discretizzate di un veicolo e produce come *output* processi sequenziali in grado di imitare, in ogni stato, la composizione parallela delle caratteristiche. Per questo obiettivo, sono necessarie le seguenti definizioni [51]:

Definizione 4.1 (Processo ben terminante): Un processo CCS è ben terminante se esegue un'azione δ se e solo se termina immediatamente dopo.

L'uso tipico di δ può essere trovato nella definizione degli operatori $\parallel$ e “;” (Tabella 13) ed è introdotto per sincronizzare i processi al termine. I due operatori corrispondono alla sequenzializzazione e all'esecuzione parallela di due processi ben terminanti. I processi risultanti dalla loro applicazione sono anch'essi ben terminanti.

La costante DONE corrisponde a un processo il cui compito è terminare senza ulteriori mosse.

L'operatore “;” verrà opportunamente usato per esprimere la sequenzialità di due processi p e q : p deve terminare prima che q possa iniziare la sua esecuzione.

Il processo “ $p||q$ ” rappresenta l'esecuzione parallela di p e q e termina solo se entrambi i processi terminano.

Nel seguito, sono considerati solo programmi CCS in cui l'operatore $||$ è stato costantemente sostituito a $|$.

$p q = (p[\delta_1/\delta] q[\delta_2/\delta] (\delta_1.\delta_2.DONE + \delta_2.\delta_1.DONE)) \setminus \{\delta_1, \delta_2\}$
$p; q = (p[\delta_3/\delta] \delta_3.q) \setminus \{\delta_3\}$
$DONE \stackrel{\text{def}}{=} \bar{\delta}.nil$

Tabella 13. Operatori per processi ben terminanti

- Costruzione del Modello

Per modellare formalmente un'immagine medica, sono utilizzate caratteristiche radiomiche, ossia dati quantitativi estratti da esami medici.

Un esame TC può essere composto da più di duecento *slices*. Da ogni *slice*, vengono estratte n caratteristiche seguendo la definizione delle caratteristiche di Pyradiomics. Per questo motivo, abbiamo estratto sette diverse classi di caratteristiche, consistenti in più di cento caratteristiche per sezione. Ogni sezione è quindi collegata in modo sequenziale per rispettare la relazione temporale e spaziale delle immagini. In questo modo, le n caratteristiche sono collegate tramite gli operatori matematici della suite dei Metodi formali. Si è utilizzato il *proc*

per definire una singola sezione che contiene la combinazione di tutte le caratteristiche estratte per una determinata classe. Si noti che questa analisi viene condotta in modo intra-classe, il che significa che ogni classe di caratteristiche produrrà un modello formale del paziente; le sezioni sono descritte attraverso la combinazione di tutte le caratteristiche di una determinata classe di caratteristiche.

Per rendere le caratteristiche indipendenti dall'ordine di estrazione, abbiamo utilizzato l'operatore a doppia barra || tra caratteristiche diverse. Una volta che le caratteristiche sono terminate, il processo richiama il processo successivo, che descrive la sezione successiva dell'esame radiologico. Quando le sezioni terminano, il processo contiene la parola nil. Ecco un esempio:

```
proc SLICEi = (F1 || F2 || ... || Fn); SLICE(i+1)
```

```
proc SLICE(i+1) = (F1 || F2 || ... || Fn); SLICE(i+k) (1)
```

```
proc SLICEk = nil
```

- Riduzione e Verifica tramite *Model checking*

Come precedentemente descritto, il modello formale consente la

creazione di tanti processi quante sono le sezioni. L'analisi viene eseguita in modalità intra-classe, quindi in ogni processo ci saranno combinazioni delle caratteristiche di una singola classe Radiomica e dei loro valori discreti all'interno di quella precisa sezione. A causa dell'esplosione degli stati, non è possibile creare un classificatore a partire da questi modelli originali. Per questo motivo, si è utilizzata una fase di riduzione al modello che consentirà di applicare la tecnica di *Model checking* e ottenere un *output* simile a una diagnosi medica.

La tecnica di riduzione viene applicata grazie alla proprietà scritta congiuntamente da informatici e specialisti medici. Essa contiene solo le informazioni rilevanti per la malattia, cioè le caratteristiche e i valori che rispettano una valutazione quantitativa della malattia. Ad esempio, si può dire che per essere malato, il paziente deve avere livelli elevati di tutte le caratteristiche di primo ordine come l'Entropia. Una volta ottenuti i modelli originali dei pazienti, e scritta la proprietà dai medici riguardo alla malattia, è possibile eseguire un'operazione di AND logico per ridurre la dimensione del modello. Infatti, ora è possibile concentrarsi solo sulle caratteristiche contenute nella proprietà, poiché queste sono quelle che guidano la decisione finale.

Questo passaggio aiuta certamente nell'identificazione delle migliori caratteristiche per l'identificazione della malattia e consente di

ridurre il modello per evitare l'esplosione degli stati. In questo modo è possibile passare alla fase di applicazione del *Model checking*, che deciderà se il modello soddisfa o meno i requisiti della proprietà.

Come anticipato, in questo studio sono state analizzate le immagini TC di 41 pazienti con diagnosi di PA moderata. Tra questi, in 16 casi vi era stata evoluzione in PA severa/necrotico-emorragica, mentre nei restanti 25 la PA moderata non era progredita.

Le features estratte e ritenute significative sono in elenco nella Tabella 14.

FIRST	Energy, Entropy, Kurtosis, Total Energy, Uniformity
GLCM	ClusterProminence, ClusterShade, Difference Average, Difference Entropy, Difference Variance
GLDM	Dependence Entropy, Dependence Variance, Large Dependence Emphasis, Large Dependence High Gray Level Emphasis, Small Dependence Emphasis

Tabella 14. Feature significative usate per definire i requisiti della proprietà

Anche in questo caso per valutare la performance di questo metodo nel classificare l'evoluzione della PA sono state usate varie metriche, tra cui *Accuracy*, *Precision*, *Recall*, *Specificity*.

Come misura dell'accuratezza è stata inoltre considerata la metrica

“*F1-Score*”, che rappresenta una media armonica di *Precision* e *Recall*, ed è particolarmente utile per valutare la performance dei modelli di classificazione, specie quando si ha a che fare con dataset poco equilibrati.

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

I risultati preliminari ottenuti mostrano una performance generale di circa l’80% (vedi Tabella 15), unitamente alla possibilità di spiegare i falsi negativi più facilmente rispetto ad altre tecniche di IA.

TP	FP	Accuracy	Precision	Recall	Specificity	F1-Score
13	5	80.5	72.2	81.3	80.0	76.5
FN	TN					
3	20					

Tabella 15. Risultati

Il modello ha quindi raggiunto un’accuratezza globale dell’80%, e solo 3 pazienti non sono stati considerati ad alto rischio di evoluzione. Questa errata diagnosi da parte del modello è stata probabilmente dovuta all’elevato grado di involuzione della ghiandola pancreatica dovuta ad altre comorbidità (es. sindrome metabolica, diabete ecc.), che ha generato una diffusa e generale riduzione delle HU a livello della ghiandola pancreatica.

CAPITOLO V

CONCLUSIONI

Il presente studio ha dimostrato che i dati di Radiomica, opportunamente analizzati mediante l'impiego dei Metodi formali, possono favorire una diagnosi automatizzata e precoce di PA moderata, garantendo un livello di attendibilità soddisfacente e un'elevata sensibilità. Tali risultati sono strettamente integrati con la supervisione e la definizione del processo da parte del medico radiologo, il cui ruolo rimane centrale per l'interpretazione e la validazione clinica dei dati.

Il contributo del medico radiologo nella strutturazione di questo strumento riveste un ruolo fondamentale nell'incrementarne l'affidabilità, garantendo che l'analisi sia conforme agli attuali standard clinici e diagnostici, nonché la più accurata e corretta possibile. Questa sinergia globale tra AI e medico radiologo contribuisce a rendere l'intero processo diagnostico altamente affidabile e clinicamente valido.

Inoltre, l'applicazione dello stesso metodo nella previsione dell'evoluzione della PA moderata in PA severa-necrotizzante ha evidenziato risultati promettenti in tal senso.

Attualmente, non sono ancora disponibili *score* clinici o radiologici in grado di prevedere l'evoluzione di una PA moderata in pancreatite necrotizzante. Pertanto, lo sviluppo di un sistema semi-automatizzato, in grado di supportare il medico radiologo nella diagnosi e nella previsione precoce dell'evoluzione della PA moderata, rappresenterebbe uno

strumento di grande rilevanza. Tale sistema consentirebbe una gestione tempestiva e accurata dei pazienti, riducendo il rischio di complicanze, migliorando gli esiti clinici e ottimizzando le risorse impiegate nella gestione di questa patologia.

Studi futuri saranno condotti con l'obiettivo di potenziare ulteriormente le prestazioni di questo strumento, estendendone l'applicazione anche a studi prospettici, al fine di consolidarne l'efficacia e favorirne l'adozione come standard innovativo nella pratica clinica.

BIBLIOGRAFIA

- [1] Phillip V, Steiner JM, Algül H. Early phase of acute pancreatitis: Assessment and management. *World J Gastrointest Pathophysiol*. 2014 Aug 15;5(3):158-68. doi: 10.4291/wjgp.v5.i3.158. PMID: 25133018; PMCID: PMC4133515.
- [2] Lee DW, Cho CM. Predicting Severity of Acute Pancreatitis. *Medicina (Kaunas)*. 2022 Jun 11;58(6):787. doi: 10.3390/medicina58060787. PMID: 35744050; PMCID: PMC9227091
- [3] Garg PK, Singh VP. Organ Failure Due to Systemic Injury in Acute Pancreatitis. *Gastroenterology*. 2019 May;156(7):2008-2023. doi: 10.1053/j.gastro.2018.12.041. Epub 2019 Feb 12. PMID: 30768987; PMCID: PMC6486861
- [4] Zilio MB, Eyff TF, Azeredo-Da-Silva ALF, Bersch VP, Osvaldt AB. A systematic review and meta-analysis of the aetiology of acute pancreatitis. *HPB (Oxford)*. 2019 Mar;21(3):259-267. doi: 10.1016/j.hpb.2018.08.003. Epub 2018 Sep 22. PMID: 30249509..
- [5] Wang GJ, Gao CF, Wei D, Wang C, Ding SQ. Acute pancreatitis: etiology and common pathogenesis. *World J Gastroenterol*. 2009 Mar 28;15(12):1427-30. doi: 10.3748/wjg.15.1427. PMID: 19322914; PMCID: PMC2665136.
- [6] Żorniak M, Beyer G, Mayerle J. Risk Stratification and Early Conservative Treatment of Acute Pancreatitis. *Visc Med*. 2019 Apr;35(2):82-89. doi: 10.1159/000497290. Epub 2019 Mar 25. PMID: 31192241; PMCID: PMC6514505.
- [7] Foster BR, Jensen KK, Bakis G, Shaaban AM, Coakley FV. Revised Atlanta Classification for Acute Pancreatitis: A Pictorial Essay. *Radiographics*. 2016 May-Jun;36(3):675-87. doi: 10.1148/rg.2016150097. Erratum in: *Radiographics*. 2019 May-Jun;39(3):912. doi: 10.1148/rg.2019194003. PMID: 27163588.
- [8] Banks PA, Bollen TL, Dervenis C, Gooszen HG, Johnson CD, Sarr MG, Tsiotos GG, Vege SS; Acute Pancreatitis Classification Working Group. Classification of acute pancreatitis--2012: revision of the Atlanta classification and definitions by international consensus. *Gut*. 2013 Jan;62(1):102-11. doi: 10.1136/gutjnl-2012-302779. Epub 2012 Oct 25. PMID: 23100216.
- [9] Glaubitz J, Asgarbeik S, Lange R, Mazloun H, Elsheikh H, Weiss FU, Sendler M. Immune response mechanisms in acute and chronic pancreatitis: strategies for therapeutic intervention. *Front Immunol*. 2023 Oct 10;14:1279539. doi: 10.3389/fimmu.2023.1279539. PMID:

37881430; PMCID: PMC10595029.

- [10] Singh VK, Bollen TL, Wu BU, Repas K, Maurer R, Yu S, Morteale KJ, Conwell DL, Banks PA. An assessment of the severity of interstitial pancreatitis. *Clin Gastroenterol Hepatol*. 2011 Dec;9(12):1098-103. doi: 10.1016/j.cgh.2011.08.026. Epub 2011 Sep 3. PMID: 21893128.
- [11] Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med*. 1995 Oct;23(10):1638-52. doi: 10.1097/00003246-199510000-00007. PMID: 7587228.
- [12] Balthazar EJ. Acute pancreatitis: assessment of severity with clinical and CT evaluation. *Radiology*. 2002 Jun;223(3):603-13. doi: 10.1148/radiol.2233010680. PMID: 12034923.
- [13] Balthazar EJ, Robinson DL, Megibow AJ, Ranson JH. Acute pancreatitis: value of CT in establishing prognosis. *Radiology*. 1990 Feb;174(2):331-6. doi: 10.1148/radiology.174.2.2296641. PMID: 2296641.
- [14] Banday IA, Gattoo I, Khan AM, Javeed J, Gupta G, Latief M. Modified Computed Tomography Severity Index for Evaluation of Acute Pancreatitis and its Correlation with Clinical Outcome: A Tertiary Care Hospital Based Observational Study. *J Clin Diagn Res*. 2015 Aug;9(8):TC01-5. doi: 10.7860/JCDR/2015/14824.6368. Epub 2015 Aug 1. PMID: 26436014; PMCID: PMC4576607.
- [15] Morteale KJ, Wiesner W, Intriere L, Shankar S, Zou KH, Kalantari BN, Perez A, vanSonnenberg E, Ros PR, Banks PA, Silverman SG. A modified CT severity index for evaluating acute pancreatitis: improved correlation with patient outcome. *AJR Am J Roentgenol*. 2004 Nov;183(5):1261-5. doi: 10.2214/ajr.183.5.1831261. PMID: 15505289.
- [16] Sahu B, Abbey P, Anand R, Kumar A, Tomer S, Malik E. Severity assessment of acute pancreatitis using CT severity index and modified CT severity index: Correlation with clinical outcomes and severity grading as per the Revised Atlanta Classification. *Indian J Radiol Imaging*. 2017 Apr-Jun;27(2):152-160. doi: 10.4103/ijri.IJRI_300_16. PMID: 28744075; PMCID: PMC5510312.
- [17] Raghuwanshi S, Gupta R, Vyas MM, Sharma R. CT Evaluation of Acute Pancreatitis and its Prognostic Correlation with CT Severity Index. *J Clin Diagn Res*. 2016 Jun;10(6):TC06-11. doi: 10.7860/JCDR/2016/19849.7934. Epub 2016 Jun 1. PMID: 27504376; PMCID: PMC4963736.

- [18] Lu JD, Cao F, Ding YX, Wu YD, Guo YL, Li F. Timing, distribution, and microbiology of infectious complications after necrotizing pancreatitis. *World J Gastroenterol*. 2019 Sep 14;25(34):5162-5173. doi: 10.3748/wjg.v25.i34.5162. PMID: 31558864; PMCID: PMC6747290
- [19] Jagielski M, Smoczyński M, Adrych K. Endoscopic treatment of walled-off pancreatic necrosis complicated with pancreaticocolonic fistula. *Surg Endosc*. 2018 Mar;32(3):1572-1580. doi: 10.1007/s00464-018-6032-4. Epub 2018 Jan 17. PMID: 29344783; PMCID: PMC5807501.
- [20] Oza VM, Kahaleh M. Endoscopic management of chronic pancreatitis. *World J Gastrointest Endosc*. 2013 Jan 16;5(1):19-28. doi: 10.4253/wjge.v5.i1.19. PMID: 23330050; PMCID: PMC3547116.
- [21] Gupta V, Krishna P, Kochhar R, Yadav TD, Bargav V, Bhalla A, Kalra N, Wig JD. Hemorrhage complicating the course of severe acute pancreatitis. *Ann Hepatobiliary Pancreat Surg*. 2020 Aug 31;24(3):292-300. doi: 10.14701/ahbps.2020.24.3.292. PMID: 32843594; PMCID: PMC7452791.
- [22] Burrowes DP, Choi HH, Rodgers SK, Fetzer DT, Kamaya A. Utility of ultrasound in acute pancreatitis. *Abdom Radiol (NY)*. 2020 May;45(5):1253-1264. doi: 10.1007/s00261-019-02364-x. PMID: 31844915.
- [23] Manikkavasakar S, AlObaidy M, Busireddy KK, Ramalho M, Nilmini V, Alagiyawanna M, Semelka RC. Magnetic resonance imaging of pancreatitis: an update. *World J Gastroenterol*. 2014 Oct 28;20(40):14760-77. doi: 10.3748/wjg.v20.i40.14760. PMID: 25356038; PMCID: PMC4209541.
- [24] Wang DB, Yu J, Fulcher AS, Turner MA. Pancreatitis in patients with pancreas divisum: imaging features at MRI and MRCP. *World J Gastroenterol*. 2013 Aug 14;19(30):4907-16. doi: 10.3748/wjg.v19.i30.4907. PMID: 23946595; PMCID: PMC3740420.
- [25] Sun H, Zuo HD, Lin Q, Yang DD, Zhou T, Tang MY, Wáng YXJ, Zhang XM. MR imaging for acute pancreatitis: the current status of clinical applications. *Ann Transl Med*. 2019 Jun;7(12):269. doi: 10.21037/atm.2019.05.37. PMID: 31355236; PMCID: PMC6614331.
- [26] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012 Mar;48(4):441-6. doi: 10.1016/j.ejca.2011.11.036. Epub 2012 Jan 16. PMID: 22257792; PMCID: PMC4533986.

- [27] Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJ, Dekker A, Fenstermacher D, Goldgof DB, Hall LO, Lambin P, Balagurunathan Y, Gatenby RA, Gillies RJ. Radiomics: the process and the challenges. *Magn Reson Imaging*. 2012 Nov;30(9):1234-48. doi: 10.1016/j.mri.2012.06.010. Epub 2012 Aug 13. PMID: 22898692; PMCID: PMC3563280.
- [28] Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016 Feb;278(2):563-77. doi: 10.1148/radiol.2015151169. Epub 2015 Nov 18. PMID: 26579733; PMCID: PMC4734157.
- [29] M. Amadasun e R. King, Textural features corresponding to textural properties, in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, n. 5, 1989, pp. 1264–1274
- [30] Depeursinge, Adrien., Al-kadi, Omar S. e Mitchell, J. Ross., *Biomedical texture analysis : fundamentals, tools 83 and challenges*, Academic Press, 2017, ISBN 9780128123218, OCLC 1002195690
- [31] Sivaramakrishna R, Powell KA, Lieber ML, Chilcote WA, Shekhar R. Texture analysis of lesions in breast ultrasound images. *Comput Med Imaging Graph*. 2002 Sep-Oct;26(5):303-7. doi: 10.1016/s0895-6111(02)00027-7. PMID: 12204234.
- [32] McGarry SD, Hurrell SL, Kaczmarowski AL, Cochran EJ, Connelly J, Rand SD, Schmainda KM, LaViolette PS. Magnetic Resonance Imaging-Based Radiomic Profiles Predict Patient Prognosis in Newly Diagnosed Glioblastoma Before Therapy. *Tomography*. 2016 Sep;2(3):223-228. doi: 10.18383/j.tom.2016.00250. PMID: 27774518; PMCID: PMC5074084
- [33] Sutton EJ, Huang EP, Drukker K, Burnside ES, Li H, Net JM, Rao A, Whitman GJ, Zuley M, Ganott M, Bonaccio E, Giger ML, Morris EA; TCGA group. Breast MRI radiomics: comparison of computer- and human-extracted imaging phenotypes. *Eur Radiol Exp*. 2017;1(1):22. doi: 10.1186/s41747-017-0025-2. Epub 2017 Nov 21. PMID: 29708200; PMCID: PMC5909355.
- [34] Cook, G.J.R., Siddique, M., Taylor, B.P. *et al*. Radiomics in PET: principles and applications. *Clin Transl Imaging* **2**, 269–276 (2014). <https://doi.org/10.1007/s40336-014-0064-0>
- [35] Parekh V, Jacobs MA. Radiomics: a new application from established techniques. *Expert Rev Precis Med Drug Dev*. 2016;1(2):207-226. doi: 10.1080/23808993.2016.1164013. Epub 2016 Mar 31. PMID: 28042608; PMCID: PMC5193485.

- [36] Edmund M. Clarke, Jr., Orna Grumberg and Doron A. Peled (2000). Model Checking
- [37] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, SMC-3(6):610–621, 1973
- [38] Kenneth I Laws. Texture energy measures. In *Proc. Image understanding workshop*, pages 47–51, 1979
- [39] Dyck, D. V. Wavelets for texture analysis, an overview. *IET Conference Proceedings* 581–585 (1997)
- [40] Susan Zannos. *Godfrey Hounsfield and the Invention of Cat Scans (Unlocking the Secrets of Science)*. Mitchell Lane Pub Inc, August 1, 2002
- [41] Mario Coriasco, Nello Balossino, Osvaldo Rampado, and Sergio Rabellino. *L'immagine digitale in diagnostica per immagini*. Springer, 2013
- [42] Mathieu Hatt, Martin Vallieres, Dimitris Visvikis, and Alex Zwanenburg. Ibsi: an international community radiomics standardization initiative, 2018
- [43] van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S, Aerts HJWL. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* 2017 Nov 1;77(21):e104-e107. doi: 10.1158/0008-5472.CAN-17-0339. PMID: 29092951; PMCID: PMC5672828
- [44] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller JV, Pieper S, Kikinis R. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging.* 2012 Nov;30(9):1323-41. doi: 10.1016/j.mri.2012.05.001. Epub 2012 Jul 6. PMID: 22770690; PMCID: PMC3466397.
- [45] Gregory M. Nielson. On marching cubes. *IEEE Transactions on visualization and computer graphics*, 9(3):283–297, 2003
- [46] Steven Salzberg. Distance metrics for instance-based learning. In *International Symposium on Methodologies for Intelligent Systems*, pages 399–408. Springer, 1991

- [47] Francesco Mercaldo, Rosangela Casolare, Giovanni Ciaramella, Giacomo Iadarola, Fabio Martinelli, Francesco Ranieri, and Antonella Santone. A realtime method for can bus intrusion detection by means of supervised machine learning. *Proceedings of the 19th 89 International Conference on Security and Cryptography (SECRYPT 2022)*, 2022
- [48] Luigi Coppolino, Roberto Nardone, Alfredo Petruolo, Luigi Romano, and Andrej Souvent. Exploiting digital twin technology for cybersecurity monitoring in smart grids. In *Proceedings of the 18th International Conference on Availability, Reliability and Security*, pages 1–10, 2023
- [49] Robin Milner. *A calculus of communicating systems*. Springer, 1980
- [50] E. M. Clarke, “Model checking,” in *Foundations of Software Technology and Theoretical Computer Science: 17th Conference Kharagpur, India, December 18–20, 1997 Proceedings* 17. Springer, 1997, pp. 54–56.
- [51] R. Milner, “Operational and algebraic semantics of concurrent processes,” in *Formal Models and Semantics*. Elsevier, 1990, pp. 1201–1242.
- [52] L. Brunese, F. Mercaldo, A. Reginelli, and A. Santone, “Prostate gleason score detection and cancer treatment through real-time formal verification,” *IEEE Access*, vol. 7, pp. 186 236–186 246, 2019. doi: 10.1109/ACCESS.2019.2961754.
- [53] A. Santone, F. Mercaldo, M. C. Brunese, F. Donnarumma, P. Guerriero, and A. Reginelli, “An innovative prostate cancer detection method via model checking,” 2021. doi: 10.21203/rs.3.rs-579730/v1.
- [54] van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S, Aerts HJWL. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* 2017 Nov 1;77(21):e104-e107. doi: 10.1158/0008-5472.CAN-17-0339. PMID: 29092951; PMCID: PMC5672828.
- [55] Eibe Frank MAH, Ian H Witten (2016) *The WEKA workbench*. Online appendix for "data mining: practical machine learning tools and techniques. Morgan Kaufmann
- [56] Pérez B (2019) Reasoning about clinical guidelines based on algebraic data types and constraint logic programming. *J Biomed Inform* 92:103134

- [57] Horak T et al (2022) Visual analysis of hyperproperties for understanding model checking results. *IEEE Trans Vis Comput Graph* 28(1):357–367
- [58] Mitchell AJ. Sensitivity $\times$ PPV is a recognized test called the clinical utility index (CUI+). *Eur J Epidemiol.* 2011 Mar;26(3):251-2; author reply 252. doi: 10.1007/s10654-011-9561-x. Epub 2011 Mar 26. PMID: 21442261.
- [59] Prajapati R, Manay P, Sugumar K, Rahandale V, Satoskar R. Acute pancreatitis: predictors of mortality, pancreatic necrosis and intervention. *Turk J Surg.* 2021 Mar 22;37(1):13-21. doi: 10.47717/turkjsurg.2021.5072. PMID: 34585089; PMCID: PMC8448565.
- [60] Neri E, Aghakhanyan G, Zerunian M, Gandolfo N, Grassi R, Miele V, Giovagnoni A, Laghi A; SIRM expert group on Artificial Intelligence. Explainable AI in radiology: a white paper of the Italian Society of Medical and Interventional Radiology. *Radiol Med.* 2023 Jun;128(6):755-764. doi: 10.1007/s11547-023-01634-5. Epub 2023 May 8. PMID: 37155000; PMCID: PMC10264482.
- [61] Kolla L, Parikh RB. Uses and limitations of artificial intelligence for oncology. *Cancer.* 2024 Jun 15;130(12):2101-2107. doi: 10.1002/cncr.35307. Epub 2024 Mar 30. PMID: 38554271; PMCID: PMC11170282.
- [62] Rocca A, Brunese MC, Santone A, Varriano G, Viganò L, Caiazzo C, Vallone G, Brunese L, Romano L, Di Serafino M; R O I Segmentation Collaborative Group. Radiomics and 256-slice-dual-energy CT in the automated diagnosis of mild acute pancreatitis: the innovation of formal methods and high-resolution CT. *Radiol Med.* 2024 Oct;129(10):1444-1453. doi: 10.1007/s11547-024-01878-9. Epub 2024 Aug 30. PMID: 39214954; PMCID: PMC11480164.
- [63] Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. *PLoS One.* 2019 Nov 7;14(11):e0224365. doi: 10.1371/journal.pone.0224365. PMID: 31697686; PMCID: PMC6837442.
- [64] Capretti G, Bonifacio C, De Palma C, Nebbia M, Giannitto C, Cancian P, Laino ME, Balzarini L, Papanikolaou N, Savevski V, Zerbi A. A machine learning risk model based on preoperative computed tomography scan to predict postoperative outcomes after pancreatoduodenectomy. *Updates Surg.* 2022 Feb;74(1):235-243. doi: 10.1007/s13304-021-01174-5. Epub 2021 Oct 1. PMID: 34596836.

- [65] Tasu JP, Guen RL, Rhouma IB, Guerrab A, Beydoun N, Bergougnot B, Ingrand P, Herpe G. Accuracy of a CT density threshold enhancement to identify pancreatic parenchyma necrosis in acute pancreatitis during the first week. *Diagn Interv Imaging*. 2022 May;103(5):266-272. doi: 10.1016/j.diii.2021.12.003. Epub 2022 Jan 3. PMID: 34991994.
- [66] Kwon Y, Park HS, Kim YJ, Jung SI, Jeon HJ. Multidetector row computed tomography of acute pancreatitis: Utility of single portal phase CT scan in short-term follow up. *Eur J Radiol*. 2012 Aug;81(8):1728-34. doi: 10.1016/j.ejrad.2011.04.061. Epub 2011 May 17. PMID: 21592705.
- [67] Park S, Chu LC, Hruban RH, Vogelstein B, Kinzler KW, Yuille AL, Fouladi DF, Shayesteh S, Ghandili S, Wolfgang CL, Burkhart R, He J, Fishman EK, Kawamoto S. Differentiating autoimmune pancreatitis from pancreatic ductal adenocarcinoma with CT radiomics features. *Diagn Interv Imaging*. 2020 Sep;101(9):555-564. doi: 10.1016/j.diii.2020.03.002. Epub 2020 Apr 8. PMID: 32278586.
- [68] Zhou Y, Ge YT, Shi XL, Wu KY, Chen WW, Ding YB, Xiao WM, Wang D, Lu GT, Hu LH. Machine learning predictive models for acute pancreatitis: A systematic review. *Int J Med Inform*. 2022 Jan;157:104641. doi: 10.1016/j.ijmedinf.2021.104641. Epub 2021 Nov 10. PMID: 34785488.
- [69] Yan G, Yan G, Li H, Liang H, Peng C, Bhetuwal A, McClure MA, Li Y, Yang G, Li Y, Zhao L, Fan X. Radiomics and Its Applications and Progress in Pancreatitis: A Current State of the Art Review. *Front Med (Lausanne)*. 2022 Jun 23;9:922299. doi: 10.3389/fmed.2022.922299. PMID: 35814756; PMCID: PMC9259974.
- [70] Zhao Y, Wei J, Xiao B, Wang L, Jiang X, Zhu Y, He W. Early prediction of acute pancreatitis severity based on changes in pancreatic and peripancreatic computed tomography radiomics nomogram. *Quant Imaging Med Surg*. 2023 Mar 1;13(3):1927-1936. doi: 10.21037/qims-22-821. Epub 2023 Feb 1. PMID: 36915340; PMCID: PMC10006146.
- [71] Sion MK, Davis KA (2019) Step-up approach for the management of pancreatic necrosis: a review of the literature. *Trauma Surg Acute Care Open* 4(1):e000308. <https://doi.org/10.1136/tsaco2019-000308>. PMID: 31245622; PMCID: PMC6560663
- [72] Mederos MA, Reber HA, Girgis MD. Acute Pancreatitis: A Review. *JAMA*. 2021 Jan 26;325(4):382-390. doi: 10.1001/jama.2020.20317. Erratum in: *JAMA*. 2021 Jun 15;325(23):2405. doi: 10.1001/jama.2021.5789. PMID: 33496779.

- [73] Ignatavicius P, Gulla A, Cernauskis K, Barauskas G, Dambrauskas Z. How severe is moderately severe acute pancreatitis? Clinical validation of revised 2012 Atlanta Classification. *World J Gastroenterol*. 2017 Nov 21;23(43):7785-7790. doi: 10.3748/wjg.v23.i43.7785. PMID: 29209119; PMCID: PMC5703938.

- [74] M. Kot, "The state explosion problem," Retrieved May, vol. 18, p. 2015, 2003

- [75] Barbuti, R., De Francesco, N., Santone, A., Vaglini, G. (1997). Selective mu-calculus: New Modal Operators for Proving Properties on Reduced Transition Systems. In: Mizuno, T., Shiratori, N., Higashino, T., Togashi, A. (eds) *Formal Description Techniques and Protocol Specification, Testing and Verification*. IFIP — The International Federation for Information Processing. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-35271-8_32

- [76] R. Barbuti, N. De Francesco, A. Santone, and G. Vaglini, "Selective mu-calculus and formula-based equivalence of transition systems," *Journal of Computer and System Sciences*, Vol. 59, No. 3, pp. 537–556, 1999.