

Alma Mater Studiorum – Università di Bologna

SCUOLA DI MEDICINA E CHIRURGIA

Corso di Laurea in Tecniche di Radiologia Medica per Immagini e Radioterapia

**L'IMPATTO DELL'INTELLIGENZA ARTIFICIALE SULLA QUALITÀ
DELLE IMMAGINI IN TC: STUDIO PILOTA PER
L'OTTIMIZZAZIONE DEI PROTOCOLLI DI ACQUISIZIONE
MEDIANTE ALGORITMI DI RICOSTRUZIONE DEEP LEARNING**

Tesi di laurea in

Laboratorio Professionale 1

Presentata da

Mattia Fontana

Relatore

Prof. Federico Tabarrini

Correlatore

Dott.ssa Silvia Strolin

Sessione I

ANNO ACCADEMICO 2022 - 2023

INDICE

CAPITOLO 1

- 1.1 - Introduzione p. 02
- 1.2 - Background p. 03
- 1.3 - Algoritmi Deep Learning p. 05

CAPITOLO 2

- 2.1 - Revisione narrativa p. 08
 - 2.1.1 - Fase 1 p. 08
 - 2.1.2 - Fase 2 p. 09
 - 2.1.3 - Fase 3 p. 10
 - 2.1.4 - Fase 4 p. 11

CAPITOLO 3

- 3.1 - Analisi di contesto p. 20
- 3.2 - Descrizione dello studio p. 21

CAPITOLO 4

- 4.1 - Analisi dei risultati p. 25

CONCLUSIONI p. 27

ALLEGATI p. 29

BIBLIOGRAFIA p. 39

RINGRAZIAMENTI p. 43

Capitolo 1

1.1 Introduzione

La Tomografia Computerizzata (TC) è una metodica utilizzata per lo studio delle componenti tissutali e vascolari del distretto addomino-pelvico che, attraverso l'uso di mezzo di contrasto, permette un confronto fra le immagini acquisite esaltandone la perfusione degli organi, il tempo di assorbimento e di eiezione dei tessuti. Nonostante sia una metodica di semplice esecuzione permette lo studio di un ampio spettro di patologie ed il continuo miglioramento delle apparecchiature nonché degli algoritmi di acquisizione o ricostruzione è spesso incentrato sull'incremento della qualità delle immagini diagnostiche, di pari passo con la riduzione dei livelli di dose radiante. Negli anni sono state avviate numerose ricerche indirizzate al confronto dei metodi di ricostruzione, dalla Filtered Back Projection (FBP) alle ricostruzioni iterative (IR) fino ai più moderni algoritmi con intelligenza artificiale (AI). Gli studi hanno incluso dataset provenienti da pazienti, fornendo di fatto una maggior quantità di informazioni sui reali valori di contrasto, sia da fantocci antropomorfi in cui l'ottimizzazione dei parametri d'acquisizione è facilitata dalla possibilità di eseguire un maggior numero di scansioni ma con informazioni meno attendibili su come approcci l'algoritmo con strutture non uniformi. Per valutare la qualità delle immagini vengono messe a confronto le acquisizioni eseguite con algoritmi di ricostruzione attraverso indagini quantitative, utilizzando per esempio il rapporto segnale-rumore (SNR) o il rapporto contrasto-rumore (CNR) come indici di qualità, o qualitative facendo valutare le immagini da Medici Specialisti in Radiologia. Ciò che emerge dalla letteratura è che gli algoritmi Deep Learning (DL) si comportano in modo più coerente a dosi decrescenti rispetto agli Algoritmi Iterativi Model-Based (MBIR) o ad alta intensità (HIR) o FBP senza compromettere la risoluzione spaziale e la rilevabilità a basso contrasto delle lesioni. Utilizzando un protocollo Deep Learning ad intensità media (DLIR-M) i valori di rumore di immagine, SNR e CNR, sono mantenuti invariati con una riduzione della dose di radiazione del 35,1% rispetto a quella del protocollo standard dose (SDCT). Nell'analisi qualitativa, la Low Dose CT con ricostruzione DLIR-M ha mostrato una migliore qualità generale dell'immagine, meno rumore e punteggi qualitativi simili o migliori per la nitidezza e la cospicuità della lesione. Il DLIR in TC addomino-pelvica a bassa

dose permette inoltre una sostanziale riduzione della dose di radiazioni, > 75%, oltre a ridurre significativamente il rumore di fondo. La sfida viene posta nel perseguire l'ottimizzazione di protocolli sul distretto addomino-pelvico agendo su parametri fisici e, nel caso di protocolli AI-Based, sull'addestramento dell'algoritmo affinché si ottenga un'elevata qualità dell'immagine diagnostica con un livello di dose adeguato che rappresenti fedelmente l'anatomia oggetto di studio.

Le radici dello studio pilota poggiano le basi su risultati in forma aggregata e/o anonima, la tesi pertanto non è stata notificata al Comitato Etico in quanto finalizzata all'acquisizione di competenze di natura metodologica per il raggiungimento di finalità didattiche.

L'elaborato valorizza, in una cornice interprofessionale, il ruolo del Tecnico Sanitario di Radiologia Medica (TSRM) .

1.2 Background

La storia della TC è caratterizzata da una rapida evoluzione delle apparecchiature a partire dal 1976, anno di sviluppo del primo scanner, in cui si faceva uso di un semplice fascio di raggi X collimato e di un singolo rivelatore che permetteva esclusivamente lo studio dell'encefalo. Il primo grande cambiamento lo si è raggiunto introducendo scanner TC che, attraverso rivelatori multipli, permettevano anche lo studio body mantenendo il movimento traslazione-rotazione delle precedenti apparecchiature. Successivamente, superato il movimento traslazionale, è stato ampliato l'array dei rivelatori permettendo il movimento sincrono con il tubo a raggi X con l'acquisizione di tutte le proiezioni necessarie alla ricostruzione delle immagini di una sezione tomografica. Il problema della velocità di acquisizione è stato risolto in un primo momento aggiungendo un meccanismo di trasferimento di segnale che utilizzava anelli slip-ring e spazzole che ha contribuito, assieme alla rotazione continua tra tubo e sistema di rivelazione, alla realizzazione della TC spirale. Attraverso l'introduzione di questa metodica si è superata l'acquisizione e la ricostruzione a singolo strato in quanto, con una sola scansione, è possibile ottenere un movimento continuo del lettino sincrono alla rotazione del sistema tubo-detettore, consentendo l'acquisizione e la ricostruzione di immagini in qualsiasi posizione all'interno del volume. La TC spirale garantisce pertanto un minor tempo di acquisizione, assenza di artefatti correlati alla mancata acquisizione di dati e ricostruzioni non assiali con una maggior copertura volumetrica. Nel tempo anche l'array dei rivelatori ha subito

un'evoluzione con l'introduzione di una matrice composta da rivelatori anche lungo la direzione "z" del lettino oltre che nel piano di rotazione del tubo, introducendo così il concetto di TC multistrato. Unitamente al banco di rivelatori si è evoluta anche la forma del fascio a raggi X, passando anche lungo la direzione "z", andando ad incidere su più file di rivelatori con relativa diminuzione del numero di rotazioni necessarie ad acquisire un volume nonché la riduzione dello spessore di strato. Questa metodica permette inoltre di ricostruire a partire da una singola acquisizione volumetrica molteplici sezioni che possono avere spessori differenti o piani differenti; negli anni l'evoluzione tecnologica ha portato la TC da 16 a 64 strati.

La Back Projection, ha sempre rappresentato un metodo di ricostruzione dell'immagine i cui valori di intensità sono distribuiti lungo la direzione della matrice; l'immagine risultante, creata dai diversi angoli di acquisizione, conteneva solo valori positivi assumendo che l'intensità del fascio fosse uniforme lungo il suo percorso, presentava una caratteristica sfocatura ai bordi. Per risolvere il problema è stato introdotto un filtro di convoluzione che filtra le proiezioni nel dominio dei dati grezzi prima di effettuare la retroproiezione, creando di fatto la Filtered Back Projection (FBP). Attraverso l'utilizzo di un kernel sarà altresì annullato l'effetto di sfocatura dell'immagine ammettendo anche valori negativi, migliorando la riproduzione dei bordi e di conseguenza la nitidezza. Nonostante questi sensibili miglioramenti la Filtered Back Projection ammette molte approssimazioni e di conseguenza è sensibile a rumori e artefatti.

Con l'aumento delle capacità computazionali sono stati introdotti metodi iterativi che oltre a ridurre la dose permettono di mantenere inalterata o migliorare la qualità dell'immagine; ciò avviene in prima battuta attraverso la creazione di dati grezzi simulati confrontati con quelli effettivamente misurati per comprendere la tipologia di correzione da applicare attraverso FBP.

I metodi iterativi principali utilizzano la statistica del conteggio dei fotoni rivelati nel processo di ricostruzione assumendo che il numero di fotoni in uscita dal tubo, e che attraversano il paziente arrivando ai detettori, abbiano una distribuzione poissoniana. Questo tipo di metodo può applicarsi nel dominio dei dati grezzi, in quello del dominio dell'immagine o in entrambi. Oppure possono esserci i metodi iterativi Model-Based che cercano di superare la statistica del conteggio dei fotoni cercando di modellizzare il processo di acquisizione nel modo più accurato possibile. Per fare ciò si utilizza la Forward

Projection (FP), in questo modo più riesco a creare un modello dell'acquisizione, più i dati grezzi simulati che si ottengono possono essere confrontati e corretti con i dati grezzi misurati.¹

1.3 Algoritmi Deep Learning

Per migliorare la qualità dell'immagine CT è stata introdotta l'intelligenza artificiale che fa uso di ricostruzioni con algoritmi Deep Learning. Il DL utilizza una rete neurale di convoluzione, quest'ultima viene addestrata creando, a partire dai dati grezzi, immagini di alta qualità. Per creare algoritmi di ricostruzione DL il primo passo è progettare l'architettura di rete impostando degli iperparametri ovvero parametri che non è possibile apprendere durante il processo di formazione (ad esempio la dimensione e la tipologia della rete).

Successivamente avviene la formazione dell'algoritmo, passaggio che richiede un set di dati costituito da più immagini di input di bassa qualità e corrispondenti immagini di alta qualità. Questi set di dati sono provenienti da scansioni con utilizzo di fantocci e da esami dei pazienti in ambito clinico; i dati utilizzati in questa fase sono suddivisi in dati di addestramento, convalida e test. Il set di dati di convalida viene utilizzato per valutare le prestazioni della rete e per ottimizzare i parametri non apprendibili che devono essere regolati manualmente prima della ricostruzione. Inoltre permette il riconoscimento dell'overfitting, ovvero riesce a comprendere quando una rete è basata sul training set per riconoscere più funzioni di quelle che forniscono i dati.² Durante il processo di addestramento supervisionato l'algoritmo DLR crea un'immagine di output partendo dai dati di input di bassa qualità che verrà immediatamente confrontata con l'immagine di alta qualità, calcolando in uscita la funzione di errore. Gli errori verranno pesati e corretti all'interno della rete, questo processo viene ripetuto in modo iterativo e la rete imparerà successivamente a eliminare la maggior parte del rumore mantenendo i dettagli anatomici. L'obiettivo delle ricostruzioni basate sul Deep Learning è di raggiungere una qualità dell'immagine simile a quella ottenuta con la FBP e con un tempo inferiore rispetto

¹ Golfieri R, Trenti R, Maccione A.G. Manuale di TC per TSRM.

² Arndt C, Güttler F, Heinrich A, Bürckenmeyer F, Diamantis I, Teichgräber U. Deep Learning CT Image Reconstruction in Clinical Practice. *Rofo*. 2021 Mar;193(3):252-261. English. doi: 10.1055/a-1248-2556. Epub 2020 Dec 10. PMID: 33302311

alle ricostruzioni iterative model-based.³ Attualmente l'utilizzo delle Convolutional Neural Networks (iologico riguarda le seguenti applicazioni: classificazione, rilevamento, segmentazione e ottimizzazione delle immagini. Un esempio riguardante la classificazione ed il rilevamento è riconoscere e caratterizzare noduli polmonari e se sono benigni o maligni. La segmentazione invece definisce i confini tra compartimenti e nell'isolare eventuali strutture, mentre per quanto riguarda l'ottimizzazione delle immagini un esempio del suo utilizzo è il denoising all'interno di una TC a basso dosaggio. I vantaggi principali nell'implementazione delle ricostruzione DL in TC riguardano la riduzione del rumore (può essere quantificato dal rapporto segnale-rumore), infatti riducendo quest'ultimo le lesioni sono più facilmente rilevabili ed in secondo luogo il rumore è direttamente associato alla dose di radiazioni. Riducendo infatti la corrente del tubo si va a ridurre la dose ma introducendo una maggiore quantità di rumore, potendo però fare affidamento su una ricostruzione DL questo aspetto può essere compensato. Ad esempio, utilizzando l'algoritmo di ricostruzione *DL Precise Image* a 5 livelli (Smoother, Smooth, Standard, Sharp, Sharper) con 5 diversi livelli di dose su dei fantocci torso e paragonandolo alle ricostruzioni FBP e una ricostruzione iterativa iDose,⁴ si è rilevato come da Smoother a Standard sia diminuita l'ampiezza del rumore per tutti i livelli di dose e lo spettro di potenza del rumore (NPS) media, mentre è aumentata la rilevabilità delle lesioni. Per quanto concerne i livelli da Standard a Sharper il pattern si è invertito; si è concluso quindi che i livelli più soft hanno ridotto il rumore e migliorato la rilevabilità delle lesioni al torace, ma hanno aumentato la levigatura dell'immagine; l'opposto è stato trovato con i livelli più nitidi. Un'altra opportunità di impiego è il miglioramento virtuale della risoluzione spaziale creando immagini con slice a spessore sottile a partire da slice a spessore più ampio. Infine un ultimo utilizzo è legato ad una migliore riduzione degli artefatti come ad esempio quelli da indurimento del fascio nell'imaging della testa e del collo. Attualmente in commercio sono disponibili diversi tipi di algoritmi di ricostruzione Deep Learning come ad esempio il Precise Image (Philips), AiCE (Canon) e il True Fidelity

³ Barca, P., Domenichelli S., Golfieri, R., Pierotti, L., Spagnoli, L., Tomasi, S., & Strigari, L. (2023). Image quality evaluation of the Precise image CT deep learning reconstruction algorithm compared to Filtered Back-projection and iDose4: a phantom study at different dose levels. *Physica Medica*, 106, 102517

⁴ Greffier J, Si-Mohamed S, Frandon J, Loisy M, de Oliveira F, Beregi JP, Dabli D. Impact of an artificial intelligence deep-learning reconstruction algorithm for CT on image quality and potential dose reduction: A phantom study. *Med Phys*. 2022 Aug;49(8):5052-5063. doi: 10.1002/mp.15807. Epub 2022 Jun 24. PMID: 35696272; PMCID: PMC9544990.

(GE Healthcare) ed è quindi fondamentale attualmente comprendere i punti di forza e quelli da migliorare in ciascuno di loro per rendere pienamente utilizzabile il loro potenziale diagnostico. Un esempio di studio comparativo è stato fatto tra AiCE e True Fidelity⁵ andando ad acquisire immagini su fantocci antropomorfi del torace a 6 diverse dosi di radiazioni andando a ricostruire i dati grezzi utilizzando la retroproiezione filtrata e i livelli DLR bassi/intermedi/più alti (L-DLR/M-DLR/H-DLR) di ciascun algoritmo. Per i livelli L-DLR/M-DLR il rumore rilevato era inferiore con TrueFidelity che con AiCE per dosi da 2,5 a 10 mGy, mentre per il livello H-DLR il rumore era inferiore con AiCE. Per L-DLR e M-DLR i valori medi della frequenza spaziale NPS erano maggiori per AiCE ad eccezione della dose 0,5 mGy. Per i livelli H-DLR l’NPS era maggiore per TrueFidelity che per AiCE. La qualità dell’immagine è stata giudicata clinicamente appropriata per tutti i livelli di entrambi gli algoritmi, per dosi da 2,5 a 10 mGy, ad eccezione del L-DLR di AiCE. Si può quindi affermare che nonostante le evidenti differenze tra diversi algoritmi si mantiene comunque un livello diagnostico accettabile andando però a diminuire il rumore rispetto ad una normale ricostruzione FBP.

⁵ Greffier J, Frandon J, Si-Mohamed S, Dabli D, Hamard A, Belaouni A, Akessoul P, Besse F, Guiu B, Beregi JP. Comparison of two deep learning image reconstruction algorithms in chest CT images: A task-based image quality assessment on phantom data. *Diagn Interv Imaging*. 2022 Jan;103(1):21-30. doi: 10.1016/j.diii.2021.08.001. Epub 2021 Sep 5. PMID: 34493475.

Capitolo 2

2.1 Revisione narrativa

La revisione della letteratura si compone di tre fasi: la prima fase è stata indirizzata sull'individuazione delle parole chiave, relative all'obiettivo di tesi, attraverso due banche dati: PubMed e Scholar. La seconda fase è stata caratterizzata dall'analisi dei titoli e degli abstract ottenuti nella fase precedente. Nella terza fase gli articoli selezionati sono stati esaminati interamente leggendo il full text ed escludendone alcuni per giungere a quelli che comporranno effettivamente la bibliografia di interesse per approcciare infine alla quarta fase, destinata unicamente a alla selezione degli articoli incentrati esclusivamente sul distretto addomino-pelvico.

2.1.1 Fase 1

La ricerca degli articoli è stata condotta nella banca dati PubMed, vedi tabella 1, in cui le parole chiave principalmente usate sono state "CT" AND "Deep Learning" AND "Image Quality" AND "Precise Image" AND "Machine Learning". Inoltre è stato anche fatto uso di altre parole chiave quali: "protocol". Per quanto concerne la ricerca sono stati utilizzati i filtri relativi alla lingua, optando per la lingua italiana e la lingua inglese, mentre da un punto di vista temporale sono state selezionate unicamente le pubblicazioni comprese nel periodo tra il 2019 e il 2023. Le combinazioni di parole chiave inserite durante la ricerca sono state: "CT protocol AND Deep Learning", "CT AND Deep Learning AND Image Quality", "Precise Image Deep Learning", "CT AND Machine Learning", "Precise Image AND Image quality", "CT protocol AND Machine Learning". Attraverso la banca dati Scholar la parola chiave che è stata utilizzata è "Precise Image" in associazione con "reconstruction" e "algorithms". Anche in questo caso i filtri relativi alla lingua includono unicamente lingua italiana e inglese, mentre per quanto riguarda il periodo di pubblicazione degli articoli sono stati presi in considerazione gli anni che vanno dal 2019 al 2023. L'unica combinazione di parole chiave che è stata utilizzata con questa banca dati è stata "Precise Image reconstruction algorithms". Durante la ricerca degli articoli è stata tenuta traccia a mezzo Excel in cui sono state segnalate la combinazione di parole chiave per la ricerca utilizzata, la banca data scelta e il numeri degli articoli risultanti in modo da poterli riportare correttamente, vedi tabella 2.

Fase 1, ricerca preliminare	
Banche dati utilizzate:	Pub Med e Scholar
Parole chiave:	CT, Deep Learning, Image Quality, Precise Image, Machine Learning.
Filtri di ricerca utilizzati:	Lingua inglese e italiana, periodo dal 2019 al 2023.
Risultato finale:	4577 articoli risultanti da analizzare.

Tabella 1

	Combinazioni di parole chiave	Articoli risultanti
1	CT protocol AND Deep Learning	8
2	CT AND Deep Learning AND Image Quality	215
3	Precise Image Deep Learning	1068
4	CT AND Machine Learning	1767
5	Precise Image AND Image quality	715
6	CT protocol AND Machine learning	9
7	"Precise Image" reconstruction algorithms	795

Tabella 2

2.1.2 Fase 2

Dopo aver quindi ottenuto i primi risultati si è passati alla Fase 2 in cui si è stato valutato il titolo di ognuno degli articoli sopra citati, allo scopo di filtrarli preliminarmente, ricavandone 1023 a partire dai 4577 iniziali.

Successivamente si è passati ad analizzare gli abstract per restringere ulteriormente il numero fonti da utilizzare in base alla coerenza e all'obiettivo; in particolare sono stati esclusi tutti quegli articoli che non riguardavano la diminuzione della dose e del rumore dell'immagine TC a parità di qualità di immagine utilizzando l'IA, sono stati esclusi anche tutti gli articoli che si concentravano principalmente su aspetti clinici e sulla segmentazione degli organi tramite IA, oltre a tutti quelli che trattavano questo argomento ma all'interno di metodiche d'acquisizione ibride come ad esempio la Cone Beam CT in Radioterapia o la PET/CT.

Il risultato finale ha portato a 37 articoli che verranno poi letti ed analizzati tramite il loro full-text in Fase 3 per determinare quali di essi saranno infine utilizzati per comporre la bibliografia.

Come per la fase 1, i dati raccolti sono stati riportati in tabella 3.

Fase 2: analisi degli articoli	
Fase preliminare:	Lettura degli articoli ed esclusione degli studi non coerenti con l'obiettivo
Risultato fase preliminare:	1023 articoli di cui analizzare l'abstract
Fase conclusiva:	Lettura degli abstract escludendo tutti quelli non riguardanti l'utilizzo dell'IA in TC per diminuire rumore e dose mantenendo il più inalterata possibile la qualità dell'immagine.
Risultato fase conclusiva:	37 studi che verranno esaminati leggendo il full-text.

Tabella 3

2.1.3 Fase 3

In questa terza fase (vedi tabella 4) sono stati esclusi tutti quei full-text riguardanti i distretti anatomici non inerenti ad addome, pelvi torace o encefalo. Sono inoltre stati eliminati studi che non riportavano al loro interno descrizioni accurate della metodologia di acquisizione dei dati e su che platea di pazienti e/o fantocci erano stati eseguiti, rimuovendo gli articoli lontani dall'approccio prettamente tecnico.

Sia durante la fase 2 che la fase 3 non sono stati considerati articoli il cui abstract o full-text non fosse disponibile.

Sono state valutate pubblicazioni che includessero l'utilizzo del deep learning confrontato con le ricostruzioni iterative o con la FBP per ridurre la dose e il rumore all'interno dell'immagine pur mantenendone la qualità, revisioni circa la storia dell'intelligenza artificiale e il suo utilizzo con il deep learning nonché revisioni sul confronto dei diversi tipi di algoritmi di ricostruzione deep learning attualmente in commercio per esaltare i punti di forza/di debolezza.

Dopo aver esaminato i 37 full-text, gli studi inclusi sono risultati 26.

Fase 3: scelta finale degli articoli	
Criteri di Inclusione:	Confronto tra algoritmi deep learning, ricostruzioni iterative e filtered back-projection cercando di diminuire la dose e rumore mantenendo la qualità dell'immagine. Revisioni riguardanti la storia delle applicazioni dell'intelligenza artificiale in TC e riguardanti gli algoritmi di deep learning attualmente in commercio.

Criteri di Esclusione:	Full text non attinenti ad addome, torace, encefalo e pelvi. Articoli non inerenti all'ambito tecnico. Articoli con metodologia e criteri di inclusione del paziente non spiegati adeguatamente.
Articoli Finali:	26 articoli compongono la bibliografia.

Tabella 4

2.1.4 Fase 4

Dai 26 full-text finali è stata eseguita un'ulteriore selezione per mettere in evidenza quali articoli fossero incentrati esclusivamente sul distretto addomino-pelvico, estraendo tutte le informazioni circa il tipo di apparecchiature utilizzate per la scansione, il tipo di algoritmo di ricostruzione, se lo studio fosse prospettico o retrospettivo e se fosse stato acquisito attraverso l'uso di un fantoccio antropomorfo.

Gli articoli a seguito dell'analisi sono risultati un totale di 9, riassunti nella tabella 5.

- “Image quality improvement with deep learning-based reconstruction on abdominal ultrahigh-resolution CT: A phantom study.” In questo studio per i modelli addominali di dimensioni normali è stato utilizzato il fantoccio Catphan 600 assieme ad un anello ovale (25 35 cm; circonferenza 95 cm). Per la valutazione dell'immagine hanno utilizzato tre moduli diversi, CTP 404, CTP 486 e CTP 515. Il fantoccio è stato scansionato utilizzando lo scanner U HRCT (Aquilion Precision; Canon Medical Systems). Le valutazioni quantitative hanno dimostrato che il DLR si comporta in modo più coerente a dosi decrescenti rispetto agli algoritmi MBIR, HIR o FBP; senza compromessi in termini di risoluzione spaziale e di rilevabilità a basso contrasto. In particolare, a dosi di radiazioni più basse, il DRL ha una valutazione quantitativa migliore del MBIR e dovrebbe ridurre il rumore dell'immagine. Per la CT dinamica addominale su U HRCT, il DLR può essere uno strumento promettente per compensare l'aumento del rumore dell'immagine dato da rilevatori più piccoli in una matrice più grande[7].
- “Image Quality and Lesion Detectability of Lower-Dose Abdominopelvic CT Obtained Using Deep Learning Image Reconstruction.” Lo scopo di questo studio non era solo quello di valutare la qualità in maniera quantitativa e qualitativa dell'immagine, ma anche di valutare la rilevabilità delle lesioni del fegato focale sulle immagini LDCT ottenute utilizzando un algoritmo DLIR rispetto a quello delle

immagini SDCT, è stato utilizzata una GE Revolution come apparecchiatura per le scansioni. Sono stati inclusi in questo studio i pazienti che hanno eseguito una TC Addomino-Pelvica con il protocollo LDCT e che ne hanno eseguita una in precedenza con il protocollo SDCT entro un anno utilizzando lo stesso scanner CT (n = 123; deviazione standard media per età [SD], 63 11 anni; maschio:femmina, 70:53). L'analisi quantitativa e qualitativa delle immagini è stata effettuata in 111 pazienti (età media SD, 63 12 anni; maschio:femmina, 62:49) dopo aver escluso dallo studio 12 pazienti che hanno subito un intervento chirurgico al fegato. Quattro pazienti che sono stati sottoposti ad ablazione a radiofrequenza sono stati inclusi nella popolazione di studio perché è stato presunto che i trattamenti locoregionali, come l'ablazione a radiofrequenza, non abbiano influenzato sia l'analisi quantitativa che qualitativa dell'immagine. La cospicuità della lesione è stata valutata in 45 dei 111 pazienti a cui è stata diagnosticata almeno una lesione focale del fegato. Le immagini LDCT sono state ricostruite con h-IR (ricostruzione statistica adattativa iterativa-Veo [ASiR-V] 40%), DLIR (TrueFidelity) a medium-strength (DLIR-M) e DLIR ad high-strength (DLIR-H), mentre le immagini SDCT sono state ricostruite con h-IR. Questo studio ha dimostrato che utilizzando un protocollo DLIR-M i valori di rumore di immagine, SNR e CNR potrebbero essere mantenuti invariati con una riduzione della dose di radiazione del 35,1% rispetto a quella del protocollo SDCT. Nell'analisi qualitativa, la LDCT con immagini DLIR-M ha mostrato una migliore qualità generale dell'immagine, meno rumore dell'immagine e punteggi qualitativi simili o migliori per la nitidezza dell'immagine e la cospicuità della lesione, anche se la trama dell'immagine non è stata la preferita. La rilevabilità delle lesioni epatiche focali su LDCT con DLIR-M non era significativamente diversa da quella di SDCT con h-IR [11].

- “Unenhanced abdominal low-dose CT reconstructed with deep learning-based image reconstruction: image quality and anatomical structure depiction. È stata valutata l'utilità dell'algoritmo di ricostruzione dell'immagine basato sull'apprendimento profondo (DLIR) nella TC addominale a basso dosaggio (LDCT). In questo studio prospettico sono stati inclusi due gruppi di pazienti: 58 pazienti sottoposti a TC addominale a dose standard senza mezzo di contrasto ricostruito

con ricostruzione iterativa ibrida (gruppo SDCT) e 48 pazienti sottoposti a LDCT addominale senza mezzo di contrasto ricostruito con un livello di intensità elevato di DLIR (gruppo LDCT) . Sono stati calcolati il rumore di fondo e il rapporto segnale/rumore (SNR) di fegato, pancreas, milza, rene, aorta addominale, vena cava inferiore e vena porta. Il rumore di fondo era inferiore nel gruppo LDCT rispetto al gruppo SDCT ($P = 0,02$). Gli SNR erano più alti nel gruppo LDCT rispetto al gruppo SDCT ($P < 0,001-0,004$) ad eccezione del fegato. Il rumore complessivo dell'immagine era superiore nel gruppo LDCT rispetto al gruppo SDCT ($P < 0,001$). La qualità complessiva dell'immagine non era diversa tra i gruppi SDCT e LDCT ($P = 0,25-0,26$). La rappresentazione di quasi tutte le strutture anatomiche addominali era uguale o migliore nel gruppo LDCT rispetto al gruppo SDCT ($P < 0,001-0,88$). L'SSDE era inferiore nel gruppo LDCT (4,0 mGy) rispetto al gruppo SDCT (20,6 mGy) ($P < 0,001$). Il DLIR facilita una sostanziale riduzione della dose di radiazioni, $> 75\%$, oltre a ridurre significativamente il rumore di fondo. Il DLIR può mantenere la qualità dell'immagine e la rappresentazione della struttura anatomica in LDCT addominale [13].

- “Impact of deep learning-based image reconstruction on image quality compared with adaptive statistical iterative reconstruction-Veo in renal and adrenal computed tomography.” E’ stata valutata la qualità dell'immagine della ricostruzione dell'immagine basata sull'apprendimento profondo (DLIR) nella tomografia computerizzata renale e surrenalica con contrasto con la ricostruzione iterativa statistica adattiva-Veo (ASiR-V). Sono stati reclutati in modo prospettico 52 pazienti. Tutte le immagini sono state ricostruite con ASiR-V 30%, ASiR-V 70% e DLIR a intensità di ricostruzione bassa, media e alta. Numero CT, rumore, tasso di riduzione del rumore, rapporto segnale-rumore (SNR) e rapporto contrasto-rumore (CNR) sono stati misurati e calcolati all'interno della regione di interesse (ROI) su grasso sottocutaneo, cortecce renali bilaterali, midollo, arterie renali e ghiandole surrenali. Per le analisi qualitative, sono state valutate la differenziazione della corteccia renale e del midollo, la visibilità del confine della ghiandola surrenale, la nitidezza, gli artefatti e il rumore soggettivo. La qualità complessiva dell'immagine è stata calcolata su una scala da 0 (peggiore) a 15

(migliore) basata sui cinque valori di cui sopra e il punteggio risultava accettabile quando ≥ 9 . Il numero CT non differisce significativamente tra i set di dati di ricostruzione. Il rumore non differisce significativamente tra ASiR-V 30% e DLIR-L, ma è significativamente inferiore utilizzando ASiR-V 70%, DLIR-M e DLIR-H. Il tasso di riduzione del rumore relativo a ASiR-V 30% è significativamente diverso tra i gruppi DLIR e ASiR-V 70% e DLIR-H produce il tasso di riduzione del rumore più elevato (61,6%). SNR e CNR sono più alti per DLIR-M, DLIR-H e ASiR-V 70% che per ASiR-V 30% e DLIR-L. DLIR-H mostra i migliori SNR e CNR. La qualità complessiva dell'immagine produce lo stesso modello per DLIR-H, con il punteggio più alto. Le percentuali di casi con punteggio complessivo della qualità dell'immagine ≥ 9 sono 100% (DLIR-H), 94,23% (DLIR-M), 90,38% (ASiR-V70%), 67,31% (DLIR-L) e 63,46% (ASiR-V30%), rispettivamente. Il DLIR ha migliorato significativamente la qualità dell'immagine oggettiva e soggettiva delle TC renali e surrenali, ottenendo una riduzione del rumore superiore rispetto a ASiR-V [14].

- “Superior objective and subjective image quality of deep learning reconstruction for low-dose abdominal CT imaging in comparison with model-based iterative reconstruction and filtered back projection.” Questo studio mirava a condurre confronti oggettivi e soggettivi della qualità dell'immagine tra le ricostruzioni di tomografia computerizzata addominale con algoritmi di ricostruzione di apprendimento profondo (DLR), ricostruzione iterativa basata su modello (MBIR) e retroproiezione filtrata (FBP). I set di dati di pazienti sono stati sottoposti a TC epatica a basse dosi e sono stati identificati retrospettivamente. Le immagini sono state ricostruite utilizzando DLR, MBIR e FBP. Sono stati calcolati il rumore medio dell'immagine e il rapporto contrasto-rumore (CNR) e il rumore, gli artefatti, la nitidezza e la qualità complessiva dell'immagine sono stati valutati soggettivamente. Sono stati inclusi novanta pazienti ($67 \pm 12,7$ anni; 63 maschi; indice medio di massa corporea [BMI], $25,5 \text{ kg/m}^2$). Il rumore medio nell'aorta addominale e nel parenchima epatico di DLR era inferiore a quello di FBP e MBIR ($p < .001$). Per FBP e MBIR, il rumore dell'immagine era significativamente più alto per i pazienti obesi rispetto a quelli con BMI normale. Il CNR per l'aorta addominale e il parenchima epatico risulta più alto per DLR che per FBP e MBIR (p

<.001). Le immagini MBIR sono state valutate soggettivamente come superiori alle immagini FBP in termini di rumore, artefatti, nitidezza e qualità complessiva ($p < .001$). Le immagini DLR sono state classificate come superiori alle immagini MBIR in termini di rumore ($p < .001$) e qualità complessiva ($p = .03$). Sulla base di confronti oggettivi e soggettivi, la qualità dell'immagine di DLR è risultata superiore a quella di MBIR e FBP su TC addominale a basso dosaggio. Il DLR era l'unico metodo per il quale il rumore dell'immagine non risultava più alto per i pazienti obesi rispetto a quelli con un BMI normale [16].

- “A study of using a deep learning image reconstruction to improve the image quality of extremely low-dose contrast-enhanced abdominal CT for patients with hepatic lesions.” Questo studio è stato eseguito per studiare la fattibilità dell'utilizzo della ricostruzione dell'immagine di apprendimento profondo (DLIR) per ridurre significativamente la dose di radiazioni e migliorare la qualità dell'immagine nella TC addominale con mezzo di contrasto, tramite uno studio prospettico. Sono stati selezionati 40 pazienti con lesioni epatiche che sono stati sottoposti a TC addominale utilizzando la dose di routine (120 kV, indice di rumore (NI) impostato su 11 con modulazione automatica della corrente del tubo) in fase arteriosa (AP) e fase portale (PP) e bassa dose (NI = 24) in fase ritardata (DP). Tutte le immagini sono state ricostruite con uno spessore di 1,25 mm utilizzando ASIR-V al 50%. Inoltre, le immagini in DP sono state ricostruite utilizzando DLIR a prestazioni elevate (DLIR-H). Sono stati misurati il valore CT e la deviazione standard (DS) del parenchima epatico, della milza, del muscolo paraspinale e della lesione. La qualità complessiva dell'immagine include il rumore soggettivo, la nitidezza, gli artefatti e l'affidabilità diagnostica sono stati valutati da due radiologi in cieco utilizzando una scala a 5 punti (1, inaccettabile e 5, eccellente). È stata confrontata la dose tra AP e DP e la qualità delle immagini tra le diverse ricostruzioni sono state comparate utilizzando SPSS 20.0. Rispetto ad AP, nella DP vi è una riduzione significativa della dose di radiazioni del 76% ($0,76 \pm 0,09$ mSv vs $3,18 \pm 0,48$ mSv), le immagini DLIR-H DP avevano un rumore dell'immagine inferiore ($14,08 \pm 2,89$ HU vs $16,67 \pm 3,74$ HU, $p < 0,001$) ma un simile punteggio complessivo della qualità dell'immagine rispetto alle immagini ASIR-V50% AP (3,88

$\pm 0,34$ vs $4,05 \pm 0,44$, $p > 0,05$). Per le immagini DP, il DLIR-H ha ridotto significativamente il rumore dell'immagine nel parenchima epatico, nella milza, nel muscolo e nella lesione a ($14,77 \pm 2,61$ HU, $14,26 \pm 2,67$ HU, $14,08 \pm 2,89$ HU e $16,25 \pm 4,42$ HU) da ($24,95 \pm 4,32$ HU, $25,42 \pm 4,99$ HU, $23,99 \pm 5,26$ HU e $27,01 \pm 7,11$) con ASIR-V50%, e punteggio di qualità dell'immagine migliorato ($3,88 \pm 0,34$ vs $2,87 \pm 0,53$; $p < 0,05$). Il DLIR-H riduce significativamente il rumore dell'immagine e genera immagini con qualità clinicamente accettabile e affidabilità diagnostica con una riduzione della dose del 76% [17].

- “Low-dose CT urography using deep learning image reconstruction: a prospective study for comparison with conventional CT urography.” Per confrontare la qualità dell'immagine dell'urografia TC a basso dosaggio (LD-CTU) è stata utilizzata la ricostruzione dell'immagine di apprendimento profondo (DLIR) con CTU convenzionale (C-CTU) confrontandolo con la ricostruzione iterativa statistica adattiva (ASIR-V). E' uno studio prospettico che utilizza le immagini CTU della fase escretoria per l'analisi. I pazienti che sono stati assegnati al gruppo LD-DLIR (100 kV e modulazione mA automatica per indice di rumore (NI) di 23) e al gruppo C-ASIR-V (100 kV e NI di 10) sono stati acquisiti secondo i protocolli di scansione nella fase escretoria. Due radiologi hanno valutato in modo indipendente la qualità complessiva dell'immagine, gli artefatti, il rumore e la nitidezza delle vie urinarie. Inoltre, sono stati valutati l'attenuazione CT media, il rapporto segnale-rumore (SNR) e il contrasto-rumore (CNR) nelle vie urinarie. Nel gruppo LD-DLIR (10 maschi e 16 femmine; età media: 57,23 anni, intervallo: 33-76 anni) e nel gruppo C-ASIR-V (14 maschi e 12 femmine; età media: 60 anni, intervallo: 33-77 anni) sono stati aggiunti 26 pazienti ciascuno. Il gruppo LD-DLIR ha utilizzato una dose di radiazioni efficace significativamente inferiore rispetto al gruppo C-ASIR-V ($2,01 \pm 0,44$ mSv vs $6,9 \pm 1,46$ mSv, $p < 0,001$). Il gruppo LD-DLIR ha mostrato una buona qualità complessiva dell'immagine con un punteggio medio >4 ed era simile a quello del gruppo C-ASIR-V. Entrambi i gruppi avevano un valore di attenuazione adeguato e simile, SNR e CNR nella maggior parte dei segmenti delle vie urinarie. È possibile quindi fornire una qualità dell'immagine comparabile riducendo al contempo la dose di radiazioni del 71% in CTU a basso dosaggio con un algoritmo

di ricostruzione dell'immagine di apprendimento profondo rispetto al CTU convenzionale con ASIR-V [20].

- “Low-Dose Abdominal CT Using a Deep Learning-Based Denoising Algorithm: A Comparison with CT Reconstructed with Filtered Back Projection or Iterative Reconstruction Algorithm.” È stata confrontata la qualità dell'immagine della tomografia computerizzata (TC) a basso dosaggio (LD) ottenuta utilizzando un algoritmo di denoising basato sull'apprendimento profondo (DLA) con immagini LD CT ricostruite con una retroproiezione filtrata (FBP) e ricostruzione iterativa modellata avanzata (ADMIRE). Cento studi TC addominali a dose di routine (RD) ricostruiti utilizzando FBP sono stati utilizzati per addestrare il DLA. Le immagini CT simulate sono state realizzate a livelli di dose del 13%, 25% e 50% della RD (DLA-1, -2 e -3) e ricostruite utilizzando FBP. Per testare il DLA, è stato utilizzato il fantoccio CT dell'American College of Radiology insieme a 18 pazienti sottoposti a TC LD addominale. Le immagini LD CT del fantoccio e dei pazienti sono state elaborate utilizzando FBP, ADMIRE e DLA (rispettivamente immagini LD-FBP, LD-ADMIRE e LD-DLA). Per confrontare la qualità dell'immagine, è stato misurato lo spettro di potenza del rumore e la funzione di trasferimento della modulazione (MTF) delle immagini fantasma. Per i dati dei pazienti, hanno misurato il rumore medio dell'immagine ed eseguito un'analisi qualitativa dell'immagine. Gli LD-DLA hanno raggiunto livelli di rumore inferiori rispetto a LD-FBP e LD-ADMIRE sia per i dati fantasma che per quelli del paziente (tutti $p < 0,001$). Gli LD-DLA addestrati con una dose di radiazioni inferiore hanno mostrato meno rumore dell'immagine. Tuttavia, gli MTF degli LD-DLA erano inferiori a quelli di LD-ADMIRE e LD-FBP (tutti $p < 0,001$) e diminuivano con la diminuzione della dose dell'immagine di addestramento. Nell'analisi qualitativa dell'immagine, la qualità complessiva degli LD-DLA era la migliore per DLA-3 (dose di radiazione simulata al 50%) e non significativamente diversa da LD-ADMIRE. Non sono stati rilevati artefatti aggiuntivi nelle immagini LD-DLA. I DLA hanno ottenuto meno rumore rispetto a FBP e ADMIRE nelle immagini LD CT, ma non hanno mantenuto la risoluzione spaziale. Il DLA addestrato con una dose di radiazioni simulata del 50% ha mostrato la migliore qualità complessiva dell'immagine [21].

- “Improved image quality and dose reduction in abdominal CT with deep-learning reconstruction algorithm: a phantom study.” E’ stato confrontato l'impatto dell'algoritmo di ricostruzione con apprendimento profondo dell'intelligenza artificiale (Precise Image; AI-DLR) sulla qualità dell'immagine rispetto a un algoritmo di ricostruzione iterativa ibrida (IR) nella TC addominale per diverse indicazioni cliniche. Le acquisizioni sui fantocci sono state eseguite a 5 livelli di dose (CTDI vol : 13/11/9/6/1,8 mGy). I dati grezzi sono stati ricostruiti utilizzando il livello 4 di iDose 4 (i4) e 3 livelli di AI-DLR (Smoother/Smooth/Standard). Sono stati calcolati lo spettro di potenza del rumore (NPS), la funzione di trasferimento basata su attività (TTF) e l'indice di rilevabilità (d'): rilevamento modellato d' di una metastasi epatica (LM) e carcinoma epatocellulare nelle fasi portale (HCCp) e arteriosa (HCCa) . La qualità dell'immagine è stata valutata soggettivamente su un fantoccio antropomorfo da 2 radiologi. Dai livelli Standard a Smoother, l'entità del rumore e la frequenza spaziale media NPS sono diminuite e la rilevabilità (d') di tutte le lesioni simulate è aumentata. Per entrambi gli inserti, i valori TTF erano simili per tutti e tre i livelli AI-DLR da 13 a 6 mGy, ma sono diminuiti dai livelli Standard a Liver a 1,8 mGy. Rispetto all'i4 utilizzato nella pratica clinica, i valori d' erano più alti utilizzando i livelli Smoother e Smooth e vicini per il livello Standard. Per tutti i livelli di dose, ad eccezione di 1,8 mGy, i radiologi hanno considerato le immagini soddisfacenti per l'uso clinico per i 3 livelli di AI-DLR, ma hanno valutato le immagini troppo uniformi utilizzando il livello Smoother. L'uso dei livelli Smooth e Smoother di AI-DLR riduce il rumore dell'immagine e migliora la rilevabilità delle lesioni e la risoluzione spaziale per livelli standard e a basso dosaggio. L'uso del livello Smooth è apparentemente il miglior compromesso tra il livello di dose più basso e un'adeguata qualità dell'immagine [23].

1	Fantoccio	DLR, MBIR, HIR o FBP	Aquilion Precision, Canon
2	Pazienti Retrospettivi	ASIR-V 40%, DLIR (True Fidelity)-M, DLIR-H	Revolution CT, GE
3	Pazienti Prospettivi	ricostruzione iterativa ibrida, DLR	Revolution CT, GE
4	Pazienti Prospettivi	ASiR-V 30%, ASiR-V 70% e DLIR a intensità di ricostruzione bassa, media e alta.	Revolution CT, GE
5	Pazienti Prospettivi	DLR, MBIR, FBP	Aquilion ONE PRISM, Canon
6	Pazienti Prospettivi	ASIR-V al 50%, DLIR-H	Revolution CT, GE
7	Pazienti Prospettivi	ASIR-V, DLIR	Revolution CT, GE
8	Fantoccio e Pazienti	DLR, FBP, ricostruzione iterativa modellata avanzata (ADMIRE).	SOMATOM Definition Edge, Siemens
9	Pazienti Prospettivi	AI-DLR (Precise Image), ricostruzione iterativa ibrida (I-Dose)	

Tabella 5

Capitolo 3

3.1 Analisi di contesto

L'U.O. Radiologia Addomino-Pelvica Diagnostica e Interventistica, si integra nel grande e complesso processo aziendale dell'IRCCS Azienda Ospedaliero-Universitaria di Bologna - Policlinico di Sant'Orsola, nella gestione della salute del paziente ricoverato ed ambulatoriale: nel primo rappresentando un anello nella catena del macroprocesso che porta dal ricovero alla dimissione, con guarigione o inquadramento diagnostico-terapeutico e nel secondo costituendo l'intero processo. In particolare, l'UO è orientata a mantenere l'elevata specializzazione nella diagnostica per immagini con moderne tecnologie US, TC e RM e la posizione di Centro di Riferimento di terzo livello Radiologia Interventistica.

In ambito TC i protocolli di acquisizione per il distretto addomino-pelvico attualmente in uso, fanno utilizzo di un sistema di ricostruzione iterativo iDose; si tratta di un algoritmo iterativo ibrido di quarta generazione introdotto da Philips Healthcare (scanner Incisive CT Philips Healthcare, Cleveland, OH, USA) dove l'elaborazione viene eseguita sia nel campo delle proiezioni che in quello dell'immagine. iDose lavora prima con i dati di proiezione, dove identifica e corregge le misurazioni TC più rumorose, successivamente, un selettore sceglie tra i modelli disponibili per la ricostruzione iterativa quello che meglio si adatta alla tipologia di volume dell'immagine.

I protocolli prevedono un'acquisizione elicoidale, con pitch di 0.80 e un tempo di rotazione di 0.50s, mentre i kV sono 120 e i mAs medi sono 115; ad essi si aggiunge un Dose Right Index di 22 con la modulazione 3D della dose. I valori fisici per questa tipologia di protocolli sono stati misurati con un Body Phantom di 32cm da cui si ricava DLP di 96.05 mGy/cm, CTDIvol di 9.38mGy un SSDE di 26.02mGy.

Il modello di utilizzo comune integra una modalità di ricostruzione iDose di livello 5 e filtro B con le seguenti caratteristiche:

- spessore di slice di 2 mm;
- incremento di 1 mm;
- matrice di 512x512px;
- finestra di visualizzazione con WC 60 e WW 360

- filtro adattivo.

Partendo da questo protocollo ne sono stati ottimizzati due analoghi con ricostruzione AI-Based (livelli standard e sharp) attraverso misurazioni mediante QRM Wire Phantom. Il fantoccio si presenta con quattro diversi tipi di densità che vanno a simulare osso, fegato, milza e polmoni così da analizzare il comportamento degli algoritmi nei vari distretti anatomici perché vengano mantenute le caratteristiche densitometriche, entro limiti accettabili per rumore e qualità di immagine.

Durante l'ottimizzazione sono stati misurati gli spessori dei bordi degli inserti per valutare la fedeltà della ricostruzione rispetto alla misura reale e per la nitidezza. Il livello *standard* AI-Based (Precise Image) utilizza una modalità di ricostruzione con i seguenti parametri:

- slice thickness di 1 mm;
- incremento di 0.5 mm;
- matrice di 512x512px
- finestra di visualizzazione con WC 60 e WW 360 e filtro adattivo.

La modalità di ricostruzione sharp mediante AI-Based si differenzia solo per il livello dell'AI impostato su *sharp*.

3.2 Descrizione dello studio

Per valutare la qualità delle immagini ricostruite mediante algoritmo Precise Image è stata redatta una griglia Excel, da sottoporre a Medico Specialista in Radiologia sulla base delle evidenze in letteratura che permette il confronto fra ricostruzioni in FBP (livello Standard), iDose (livello 2, 4, 5, 7) e i livelli Precise Image (livello Smoother, Smooth, Standard).

Al contrario dell'acquisizione sul fantoccio, la cui raccolta di dati fisici permette di studiare la fedeltà della ricostruzione alla realtà del volume acquisito, la valutazione iconografica da parte di un osservatore su immagini reali, permette di avere informazioni sulla qualità del risultato nonché sulla congruenza rispetto al quesito diagnostico posto.

Per l'impostazione della griglia excel sono state prese in considerazione le principali evidenze in letteratura con focus al confronto della ricostruzione AI-Based con FBP e ricostruzioni iterative.

- Studio 1: è stata eseguita una TC torace su un fantoccio i cui raw data sono stati rielaborati in FBP filter B, iDose Level 4/7 Filter B e Precise Image in 5 livelli (Smoother, Smooth, Standard, Sharp, Sharper). Le immagini risultanti sono state

sottoposte all'osservazione di due Specialisti in Radiologia in cieco (senza conoscere la tipologia di ricostruzione ne dose level) valutando, mediante scala Likert da 1 a 5 punti, il rumore e la "morbidezza" dell'immagine, mentre con scala da 1 a 4 il contrasto tra vasi e grasso nel mediastino ed il contrasto tra bronchi e parenchima polmonare; un valore inferiore a 3 è considerato insoddisfacente a livello clinico.⁶

- Studio 2: è stata eseguita una TC colonna vertebrale su fantoccio; in questo caso i raw data sono stati rielaborati con algoritmi iDose 4 e Precise Image (Smoother, Smooth e Standard). Il distretto è stato valutato da uno Specialista in Radiologia senior ed uno junior attraverso una scala da 1 a 4 con focus sul percepito in termini di qualità generale delle immagini in "morbidezza" e rumore⁷, dove valori inferiori di 3 sono considerati insufficienti.
- Studio 3: è stata eseguita una TC addome su 30 pazienti con almeno una metastasi al fegato conosciuta; i raw data sono stati ricostruiti con iDose 4 e Precise Image (Smoother, Smooth e Standard). Le immagini sono state successivamente valutate da due Specialisti in Radiologia tramite scala di Likert da 1 a 5 con focus su rumore e "morbidezza", mentre con scala da 1 a 4 la qualità generale dell'immagine.⁸
- Studio 4: è stata eseguita una TC torace su fantoccio con raw data rielaborati con FBP, e due diversi tipi di algoritmi AI-Based (TrueFidelity e AiCE) con livello lowest/intermediate/highest. Le immagini sono state sottoposte all'osservazione di due Specialisti in Radiologia senior che hanno valutato mediante scala Likert da 1 a 5 il rumore, la "morbidezza" e il contrasto tra vasi e grasso, mentre tramite scala da 1 a 4 è stata valutata la qualità generale.⁹

⁶ Greffier J, Si-Mohamed S, Frandon J, Loisy M, de Oliveira F, Beregi JP, Dabli D. Impact of an artificial intelligence deep-learning reconstruction algorithm for CT on image quality and potential dose reduction: A phantom study. *Med Phys.* 2022 Aug;49(8):5052-5063. doi: 10.1002/mp.15807. Epub 2022 Jun 24. PMID: 35696272; PMCID: PMC9544990.

⁷ Greffier J, Frandon J, Durand Q, Kammoun T, Loisy M, Beregi JP, Dabli D. Contribution of an artificial intelligence deep-learning reconstruction algorithm for dose optimization in lumbar spine CT examination: A phantom study. *Diagn Interv Imaging.* 2023 Feb;104(2):76-83. doi: 10.1016/j.diii.2022.08.004. Epub 2022 Sep 11. PMID: 36100524.

⁸ Greffier J, Durand Q, Serrand C, Sales R, de Oliveira F, Beregi JP, Dabli D, Frandon J. First Results of a New Deep Learning Reconstruction Algorithm on Image Quality and Liver Metastasis Conspicuity for Abdominal Low-Dose CT. *Diagnostics (Basel).* 2023 Mar 20;13(6):1182. doi: 10.3390/diagnostics13061182. PMID: 36980490; PMCID: PMC10047497.

⁹ Greffier J, Frandon J, Si-Mohamed S, Dabli D, Hamard A, Belaouni A, Akessoul P, Besse F, Guiu B, Beregi JP. Comparison of two deep learning image reconstruction algorithms in chest CT images: A task-based image

Images Key	Rumore (1 Liscio, 3 Standard, 5 Granuloso)					Artefatto (Yes/No)	AI Level - Livello di Soddisfazione							
	1	2	3	4	5		1	2	3	4	5			
Std														
iD_2													1	Insoddisfacente
iD_4													2	Scarsa
iD_5													3	Discreto
iD_7													4	Buono
P_Sr													5	Molto buono
P_S														
P_St														

Tabella 7

Study ID	ID Patient	Data CT	Sesso	Età	Radiologo (I/S)	Distretto	Recon Mode	Filter	Level	Images Key	Rumore (1 Liscio, 3 Standard, 5 Granuloso)					Artefatto (Yes/No)	AI Level - Livello di Soddisfazione				
											1	2	3	4	5		1	2	3	4	5
11	124816	20/9	F	49		Rene	Standard (FBP)	B	-	Std											
							iDose	B	2	iD_2											
							iDose	B	4	iD_4											
							iDose	B	5	iD_5											
							iDose	B	7	iD_7											
							Precise Image	B	Smoother	P_Sr											
							Precise Image	B	Smooth	P_S											
							Precise Image	B	Standard	P_St											

Tabella 8

Capitolo 4

4.1 Analisi dei risultati

I risultati emersi, sono stati interpretati attraverso l'utilizzo di grafici statistici violin plot per la valutazione del rumore (vedi tabella 8) che sulla qualità generale dell'immagine (vedi tabella 9).

Ogni grafico rappresenta una tipologia di ricostruzione diversa tra FBP, iDose e Precise Image. Dalla valutazione del rumore nel fegato emerge come nell'iDose, Level 2 e 4, la trama dell'immagine risulti prevalentemente granulosa mentre nel Level 5 sia considerata ad un grado standard. L'iDose level 7 appare invece molto simile al Precise Image Standard risultando più omogeneo dei precedenti livelli. Il Precise Image (Smooth e Smoother) risulta omogeneo mentre l'FBP viene valutato come molto granuloso.

Prendendo in considerazione il distretto renale i risultati risultano più polarizzati mantenendo però le stesse caratteristiche del distretto epatico con l'iDose level 2 e FBP che appaiono più granulose, iDose level 4 e level 5 così come il Precise Image raggiungono un punteggio di 3 ed infine il Precise Image Smooth e Smoother risultano le più omogenee.

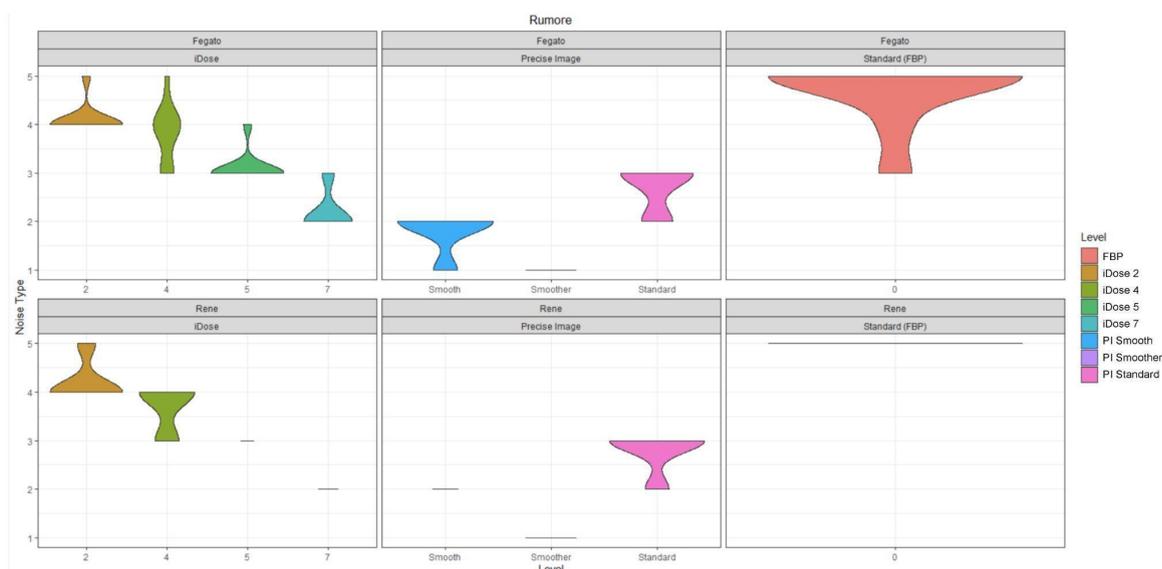


Tabella 8

Nella valutazione del gradimento dell'immagine è stato riscontrato come nel fegato la ricostruzione con i valori più alti sia il Precise Image Standard seguito dall'iDose Level 4 e

Level 5, mentre l'iDose Level 2 e Level 7 hanno raggiunto una valutazione tra il tre ed il quattro (così come il Precise Image Smoother e Smooth che si attestano come ricostruzioni valide). Il FBP invece rappresenta l'alternativa con la peggior valutazione tra le otto prese in considerazione all'interno di questo studio.

Il rene mostra una situazione leggermente diversa in cui il Precise Image Standard risulta il migliore e abbastanza in linea con l'iDose 4 e 5; l'iDose 2 risulta essere discreto così come il Precise Image Smooth; le ricostruzioni meno soddisfacenti risultano essere il Precise Image Smoother e l'FBP, mentre l'iDose 7 mostra valutazioni negative non risultando pienamente affidabile.

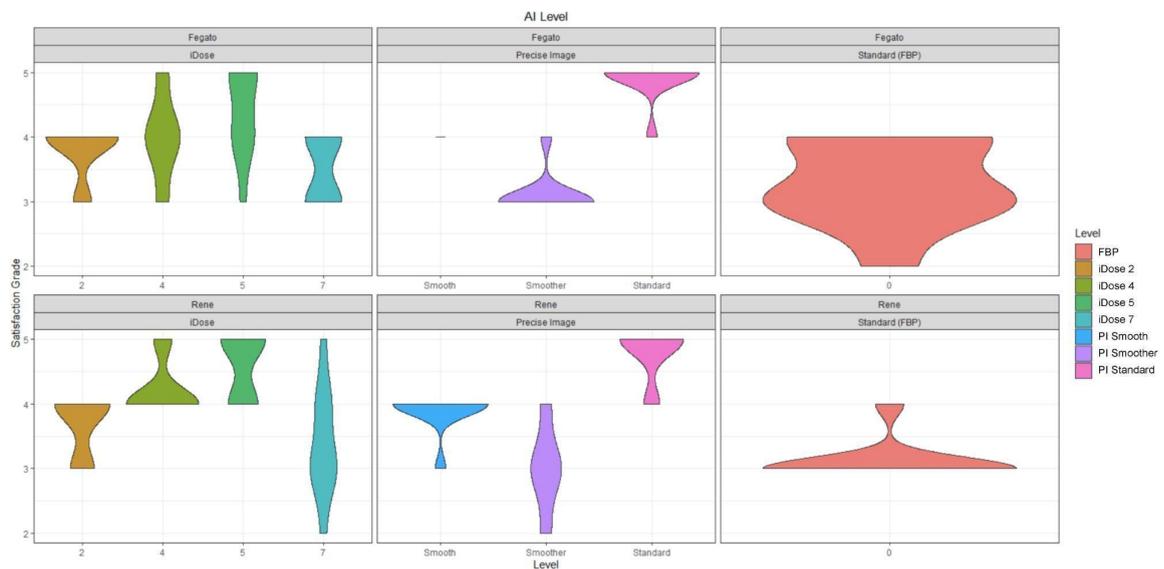


Tabella 9

Conclusioni

L'opportunità di contribuire, come laureando, al miglioramento di un servizio sanitario nell'ambito di un progetto di ottimizzazione, offre ampi spunti riflessivi su un'attività professionale che nel corso di questi ultimi anni ha richiesto maggiori competenze, in termini di valore, qualità, conoscenze, doti, capacità ed abilità.

I risultati dell'elaborato, portano ad importanti considerazioni collegate direttamente all'esperienza del singolo osservatore coinvolto (Medico Specialista in Radiologia), nonché ai potenziali input da valutare in uno studio finalizzato all'ottimizzazione di protocolli TC mediante l'utilizzo di Intelligenza Artificiale. L'analisi delle fonti bibliografiche evidenzia la necessità di condurre ulteriori studi basati sul confronto tra diverse modalità di ricostruzione, utilizzando sia valutazioni quantitative che qualitative da parte di osservatori specializzati, su un campione significativamente più ampio di pazienti. Molti degli studi citati in bibliografia si concentrano su acquisizioni effettuate su fantocci o su piccoli gruppi di pazienti; tuttavia, coinvolgere un campione di pazienti più ampio (ad esempio 40-50 per ciascun distretto) consentirebbe di ottenere valutazioni statisticamente più significative nonché di associare in modo accurato le diverse tecniche di ricostruzione ai distretti anatomici di interesse. L'analisi dei risultati suggerisce, infine, la necessità di utilizzare algoritmi diversi in base all'oggetto di studio; ne è un esempio il *Precise Image livello Standard* risultato particolarmente adattabile alle indicazioni cliniche nello studio del distretto epatico, con valutazioni elevate sia sulla qualità generale delle immagini che sul rumore. Tuttavia, questa accuratezza varia quando si approccia ad altre strutture anatomiche, quali il rene, dove la qualità mostrata risulta simile ai risultati iconografici attraverso l'uso di *iDose livello 4 o 5* ma con trama più "morbida".

Per ottimizzare i protocolli di acquisizione/ricostruzione oggetto dell'elaborato, appare ipotizzabile adattare regioni di interesse (ROI) più specifiche, in modo da comprendere quale algoritmo risulti più efficace nell'individuazione di un tessuto o una patologia. Questa differenziazione consentirebbe una personalizzazione congruente alle indagini fornendo agli specialisti immagini di alta qualità per una diagnosi più accurata.

Non ultimo la valutazione delle immagini da parte di un gruppo più ampio di Medici Specialisti con diversi gradi di expertise orientati alla diagnostica per immagini TC in radiologia addomino-pelvica rafforzerebbe la validità statistica del campione.

Il valore aggiunto della duplice valutazione inter-operatore e intra-operatore consentirebbe infine di estrapolare informazioni con una discreta variabilità tra osservatore Junior e Senior, dovuta inevitabilmente agli anni di esperienza e di competenze acquisite sul campo.

L'implementazione di questi accorgimenti potrebbe contribuire lo sviluppo di protocolli più accurati e personalizzati, migliorando di fatto la diagnostica radiologica, la qualità dell'assistenza nonché riducendo ulteriormente il livello di dose assorbito dal paziente grazie all'impiego di algoritmi AI rispetto a quelli iterativi di routine clinica.

Allegati

Di seguito sono riportate le immagini relative ad uno dei casi clinici, oggetto di studio, utilizzate per permettere allo Specialista in Radiologia di valutare rumore e qualità delle immagini. Partendo dalla TC Basale (immagini 1 e 10), sono state inserite in sequenza la ricostruzione FBP Standard, iDose Level 2/4/5/7, Precise Image Level Smoother/Smooth/Standard. Per rendere più semplice l'identificazione le ricostruzioni sono state associati specifici Label Key: Basale A.I, Std (immagini 2 e 11), iD_2 (immagini 3 e 12), iD_4 (immagini 4 e 13), iD_5 (immagini 5 e 14), iD_7 (immagini 6 e 15), P_Sr (immagini 7 e 16), P_S (immagini 8 e 17), P_St (immagini 9 e 18).

Distretto Epatico



Immagine 1

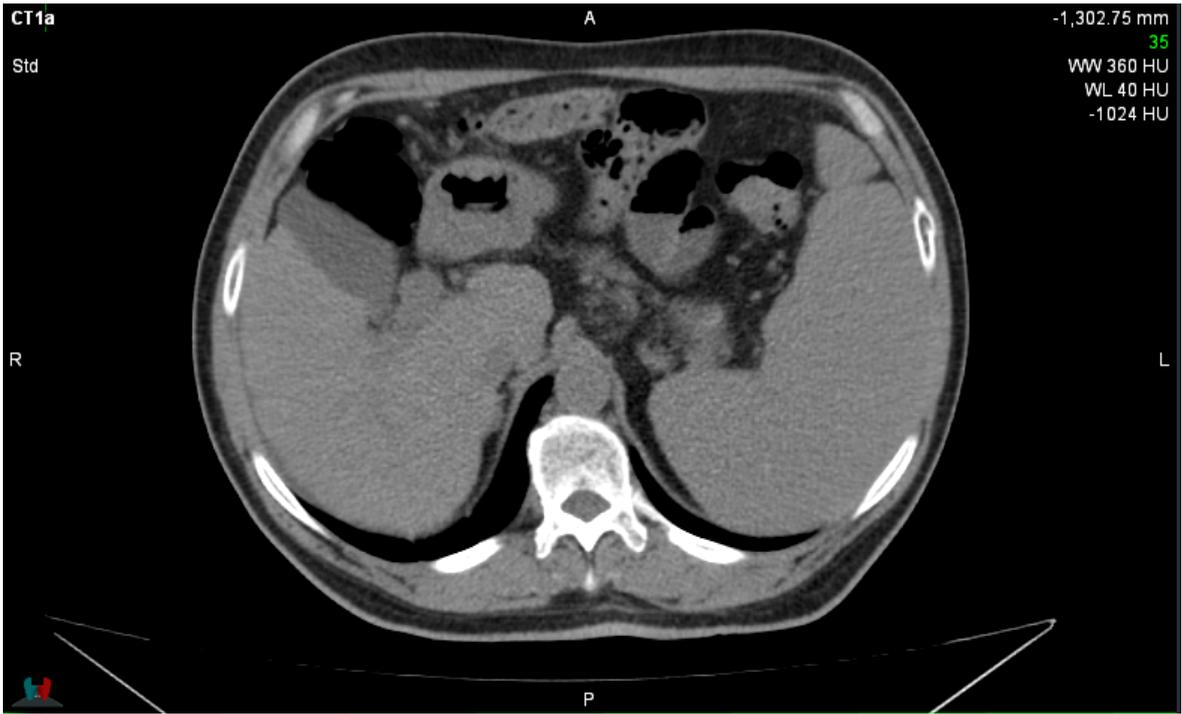


Immagine 2

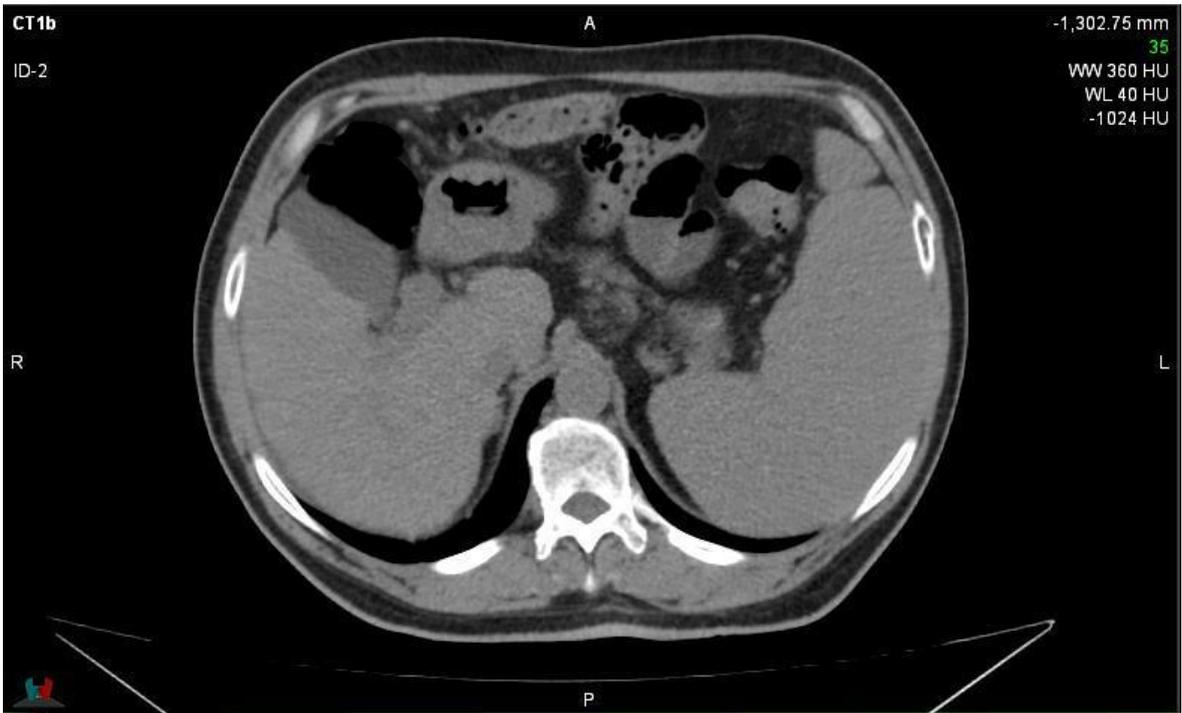


Immagine 3

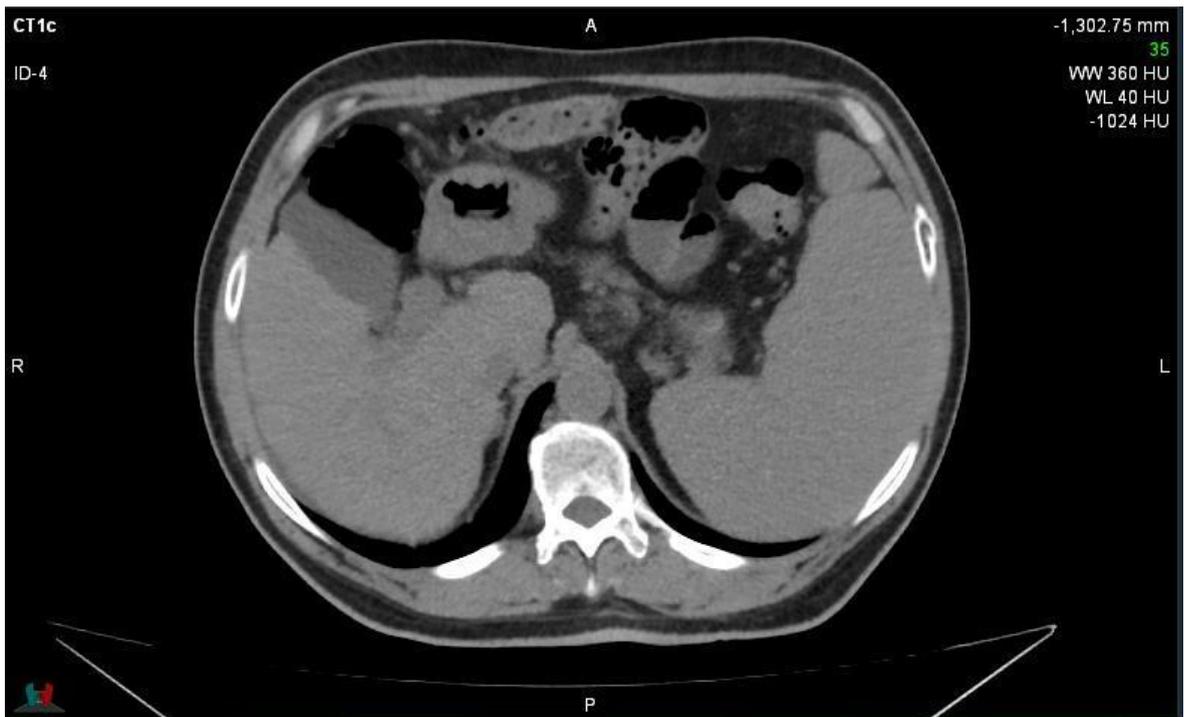


Immagine 4

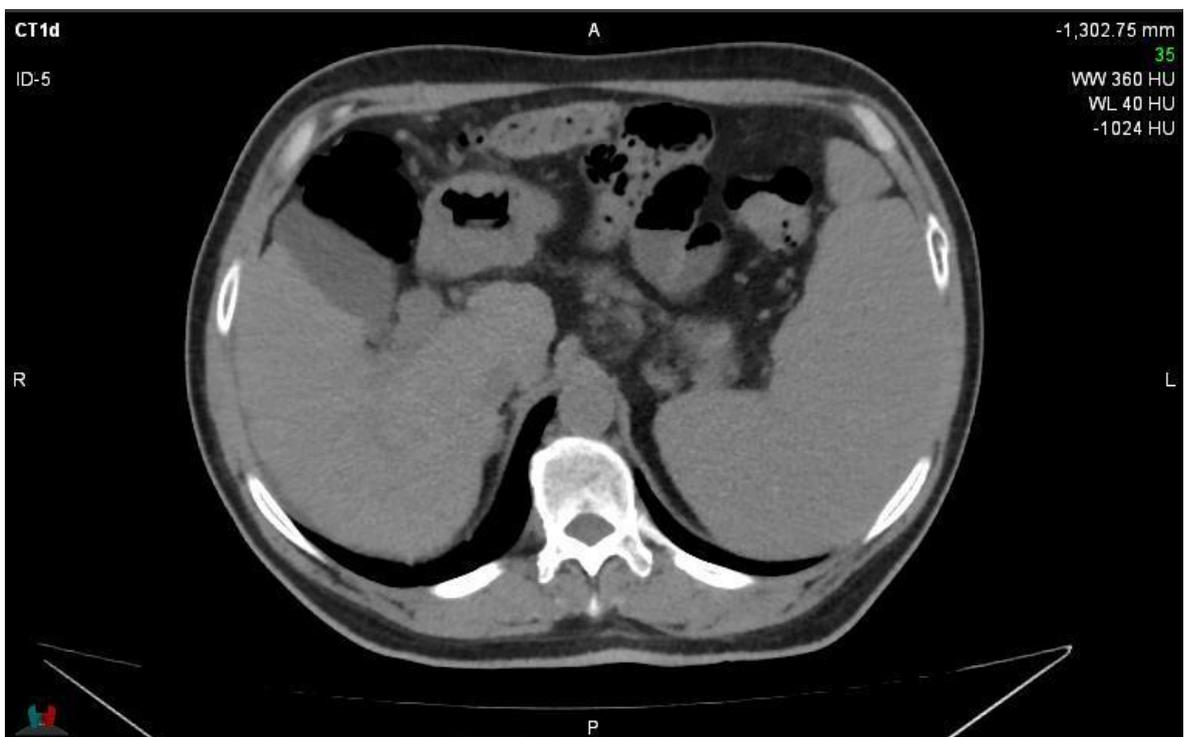


Immagine 5

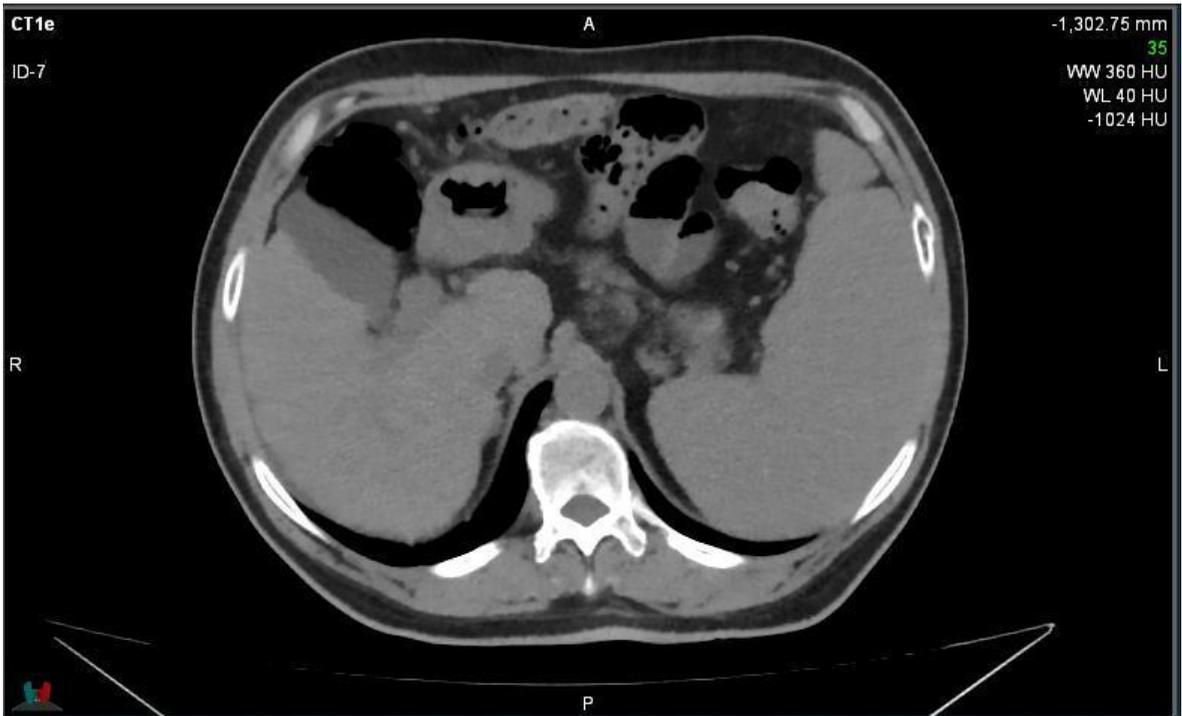


Immagine 6

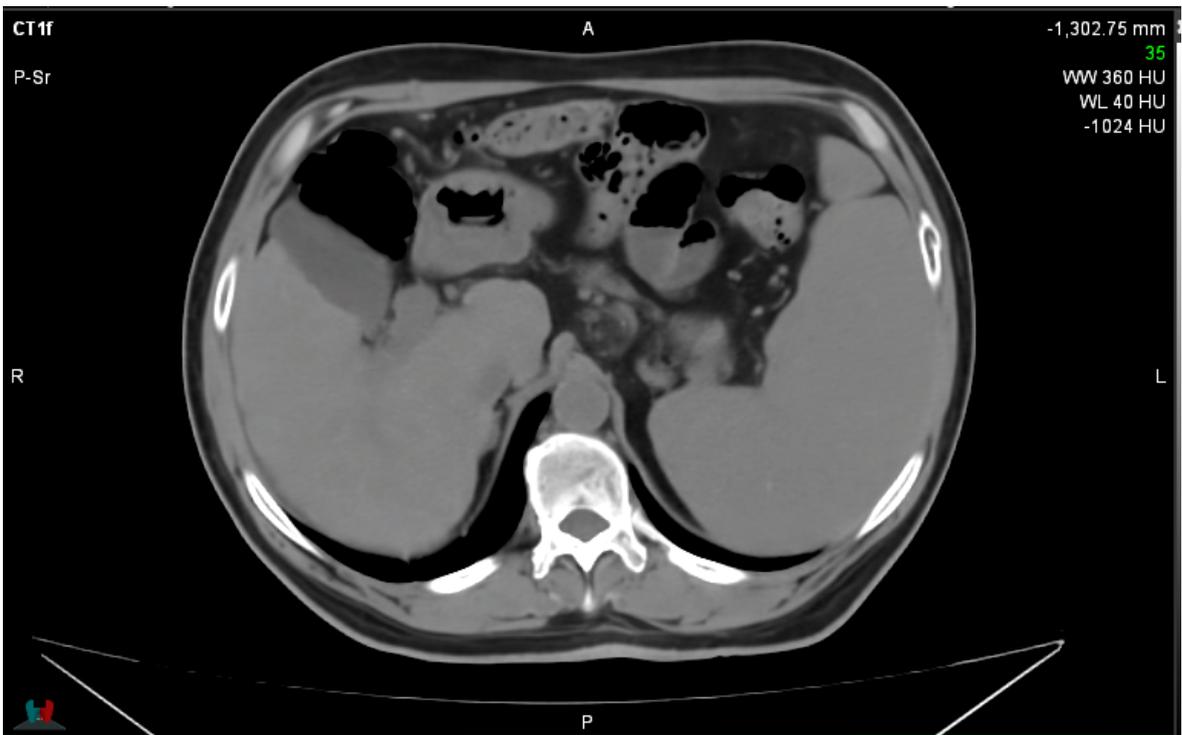


Immagine 7

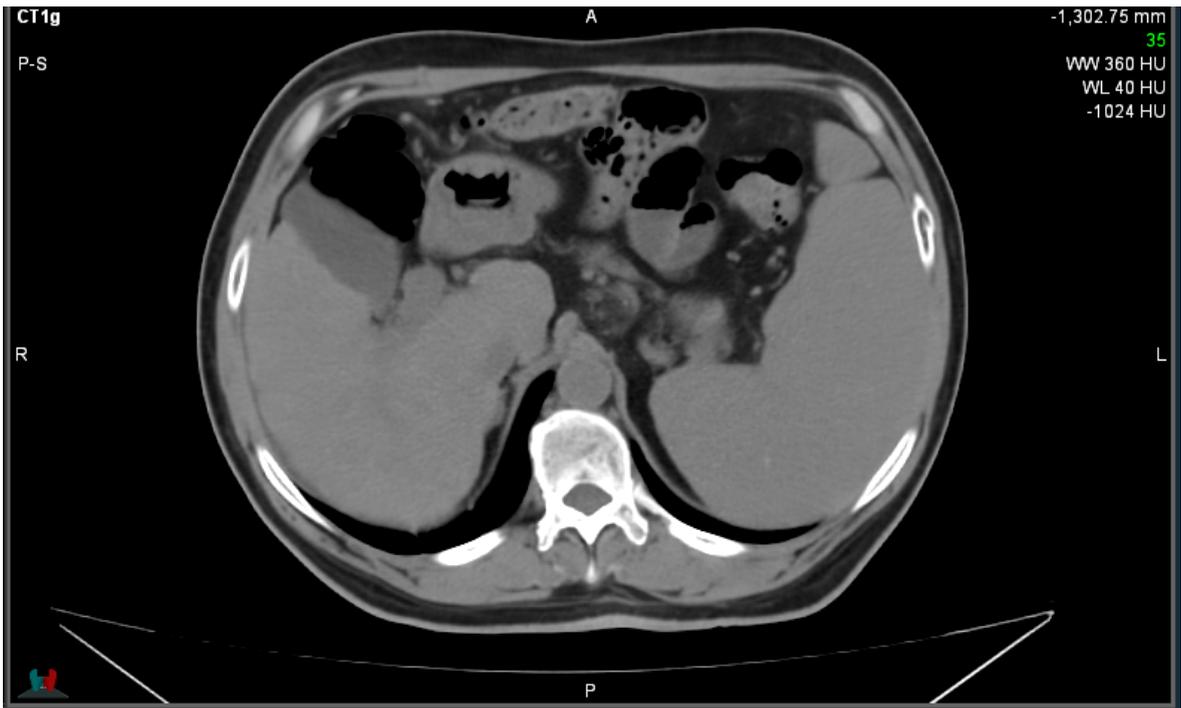


Immagine 8

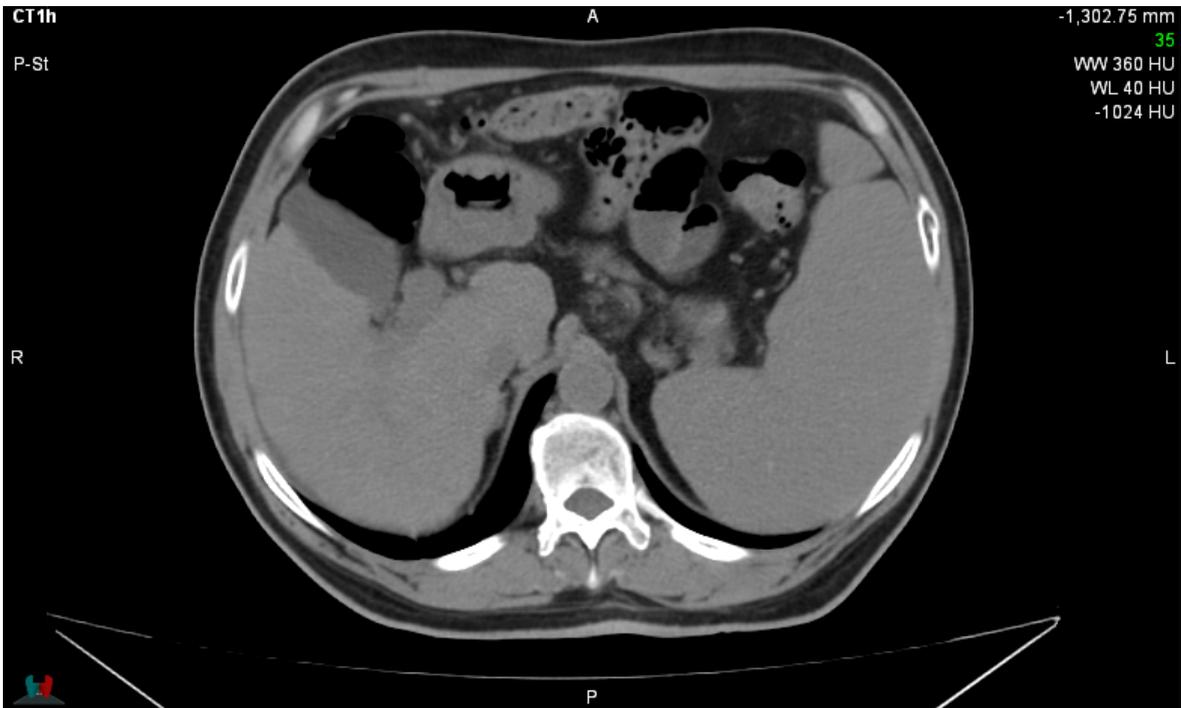


Immagine 9

Distretto Renale



Immagine 10



Immagine 11



Immagine 12



Immagine 13



Immagine 14



Immagine 15

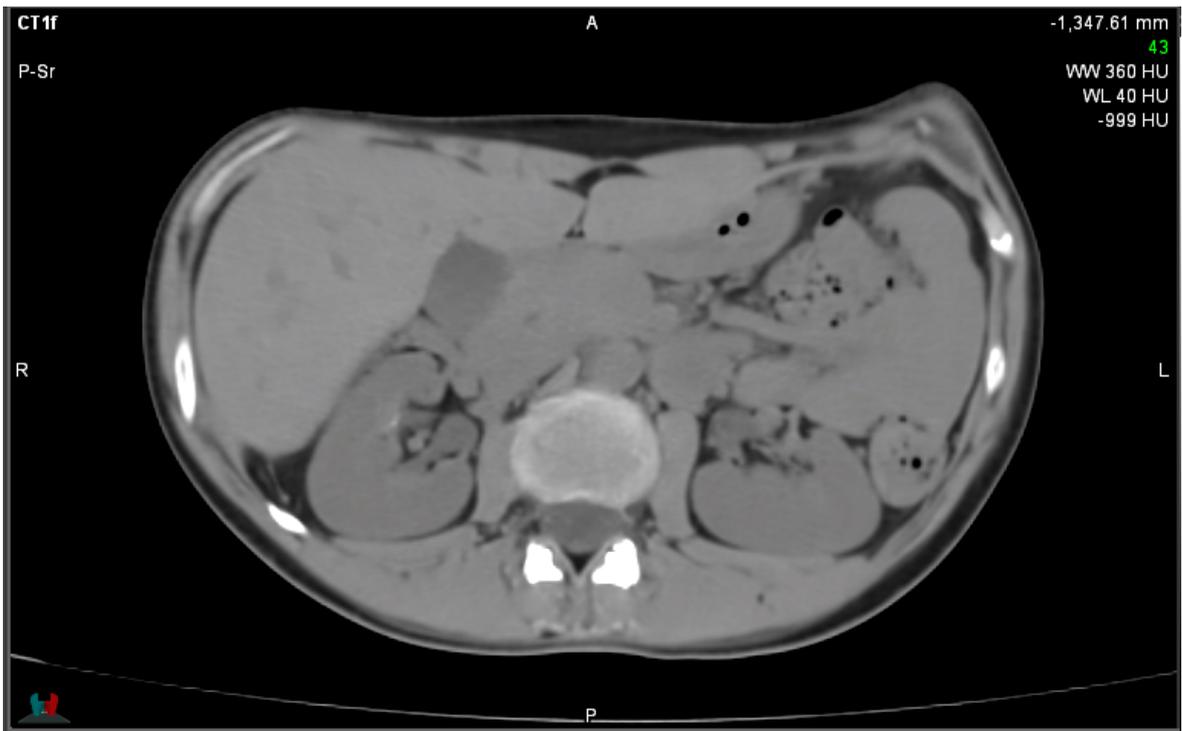


Immagine 16



Immagine 17



Immagine 18

Bibliografia

1. Golfieri R, Trenti R, Maccione A.G. Manuale di TC per TSRM.
2. Arndt C, Güttler F, Heinrich A, Bürckenmeyer F, Diamantis I, Teichgräber U. Deep Learning CT Image Reconstruction in Clinical Practice. *Rofo*. 2021 Mar;193(3):252-261. English. doi: 10.1055/a-1248-2556. Epub 2020 Dec 10. PMID: 33302311.
3. Barca, P., Domenichelli S., Golfieri, R., Pierotti, L., Spagnoli, L., Tomasi, S., & Strigari, L. (2023). Image quality evaluation of the Precise image CT deep learning reconstruction algorithm compared to Filtered Back-projection and iDose4: a phantom study at different dose levels. *Physica Medica*, 106, 102517.
4. Greffier J, Si-Mohamed S, Frandon J, Loisy M, de Oliveira F, Beregi JP, Dabli D. Impact of an artificial intelligence deep-learning reconstruction algorithm for CT on image quality and potential dose reduction: A phantom study. *Med Phys*. 2022 Aug;49(8):5052-5063. doi: 10.1002/mp.15807. Epub 2022 Jun 24. PMID: 35696272; PMCID: PMC9544990.
5. Greffier J, Frandon J, Si-Mohamed S, Dabli D, Hamard A, Belaouni A, Akessoul P, Besse F, Guiu B, Beregi JP. Comparison of two deep learning image reconstruction algorithms in chest CT images: A task-based image quality assessment on phantom data. *Diagn Interv Imaging*. 2022 Jan;103(1):21-30. doi: 10.1016/j.diii.2021.08.001. Epub 2021 Sep 5. PMID: 34493475.
6. Salimi Y, Shiri I, Akhavanallaf A, Mansouri Z, Sanaat A, Pakbin M, Ghasemian M, Arabi H, Zaidi H. Deep Learning-based calculation of patient size and attenuation surrogates from localizer Image: Toward personalized chest CT protocol optimization. *Eur J Radiol*. 2022 Dec;157:110602. doi: 10.1016/j.ejrad.2022.110602. Epub 2022 Nov 11. PMID: 36410091.
7. Shirasaka T, Kojima T, Funama Y, Sakai Y, Kondo M, Mikayama R, Hamasaki H, Kato T, Ushijima Y, Asayama Y, Nishie A. Image quality improvement with deep learning-based reconstruction on abdominal ultrahigh-resolution CT: A phantom study. *J Appl Clin Med Phys*. 2021 Jul;22(7):286-296. doi: 10.1002/acm2.13318. Epub 2021 Jun 23. PMID: 34159736; PMCID: PMC8292685.
8. van Stiphout JA, Driessen J, Koetzier LR, Ruules LB, Willemink MJ, Heemskerk JWT, van der Molen AJ. The effect of deep learning reconstruction on abdominal CT

- densitometry and image quality: a systematic review and meta-analysis. *Eur Radiol.* 2022 May;32(5):2921-2929. doi: 10.1007/s00330-021-08438-z. Epub 2021 Dec 15. PMID: 34913104; PMCID: PMC9038933.
9. Tian Q, Li X, Li J, Cheng Y, Niu X, Zhu S, Xu W, Guo J. Image quality improvement in low-dose chest CT with deep learning image reconstruction. *J Appl Clin Med Phys.* 2022 Dec;23(12):e13796. doi: 10.1002/acm2.13796. Epub 2022 Oct 9. PMID: 36210060; PMCID: PMC9797160.
 10. Yoon H, Kim J, Lim HJ, Lee MJ. Image quality assessment of pediatric chest and abdomen CT by deep learning reconstruction. *BMC Med Imaging.* 2021 Oct 10;21(1):146. doi: 10.1186/s12880-021-00677-2. PMID: 34629049; PMCID: PMC8503996.
 11. Park J, Shin J, Min IK, Bae H, Kim YE, Chung YE. Image Quality and Lesion Detectability of Lower-Dose Abdominopelvic CT Obtained Using Deep Learning Image Reconstruction. *Korean J Radiol.* 2022 Apr;23(4):402-412. doi: 10.3348/kjr.2021.0683. Epub 2022 Jan 27. PMID: 35289146; PMCID: PMC8961013.
 12. Noda Y, Kaga T, Kawai N, Miyoshi T, Kawada H, Hyodo F, Kambadakone A, Matsuo M. Low-dose whole-body CT using deep learning image reconstruction: image quality and lesion detection. *Br J Radiol.* 2021 May 1;94(1121):20201329. doi: 10.1259/bjr.20201329. Epub 2021 Feb 22. PMID: 33571010; PMCID: PMC8506192.
 13. Kaga T, Noda Y, Mori T, Kawai N, Miyoshi T, Hyodo F, Kato H, Matsuo M. Unenhanced abdominal low-dose CT reconstructed with deep learning-based image reconstruction: image quality and anatomical structure depiction. *Jpn J Radiol.* 2022 Jul;40(7):703-711. doi: 10.1007/s11604-022-01259-0. Epub 2022 Mar 14. PMID: 35286578; PMCID: PMC9252942.
 14. Bie Y, Yang S, Li X, Zhao K, Zhang C, Zhong H. Impact of deep learning-based image reconstruction on image quality compared with adaptive statistical iterative reconstruction-Veo in renal and adrenal computed tomography. *J Xray Sci Technol.* 2022;30(3):409-418. doi: 10.3233/XST-211105. PMID: 35124575; PMCID: PMC9108564.
 15. Kim JH, Yoon HJ, Lee E, Kim I, Cha YK, Bak SH. Validation of Deep-Learning Image Reconstruction for Low-Dose Chest Computed Tomography Scan: Emphasis on Image

- Quality and Noise. *Korean J Radiol.* 2021 Jan;22(1):131-138. doi: 10.3348/kjr.2020.0116. Epub 2020 Jul 27. PMID: 32729277; PMCID: PMC7772377.
16. Tamura A, Mukaida E, Ota Y, Kamata M, Abe S, Yoshioka K. Superior objective and subjective image quality of deep learning reconstruction for low-dose abdominal CT imaging in comparison with model-based iterative reconstruction and filtered back projection. *Br J Radiol.* 2021 Jul 1;94(1123):20201357. doi: 10.1259/bjr.20201357. PMID: 34142867; PMCID: PMC8248220.
 17. Cao L, Liu X, Li J, Qu T, Chen L, Cheng Y, Hu J, Sun J, Guo J. A study of using a deep learning image reconstruction to improve the image quality of extremely low-dose contrast-enhanced abdominal CT for patients with hepatic lesions. *Br J Radiol.* 2021 Feb 1;94(1118):20201086. doi: 10.1259/bjr.20201086. Epub 2020 Dec 11. PMID: 33242256; PMCID: PMC7934287.
 18. Sun J, Li H, Wang B, Li J, Li M, Zhou Z, Peng Y. Application of a deep learning image reconstruction (DLIR) algorithm in head CT imaging for children to improve image quality and lesion detection. *BMC Med Imaging.* 2021 Jul 8;21(1):108. doi: 10.1186/s12880-021-00637-w. PMID: 34238229; PMCID: PMC8268450.
 19. Son W, Kim M, Hwang JY, Kim YW, Park C, Choo KS, Kim TU, Jang JY. Comparison of a Deep Learning-Based Reconstruction Algorithm with Filtered Back Projection and Iterative Reconstruction Algorithms for Pediatric Abdominopelvic CT. *Korean J Radiol.* 2022 Jul;23(7):752-762. doi: 10.3348/kjr.2021.0466. Epub 2022 May 27. PMID: 35695313; PMCID: PMC9240291.
 20. Cheng Y, Han Y, Li J, Fan G, Cao L, Li J, Jia X, Yang J, Guo J. Low-dose CT urography using deep learning image reconstruction: a prospective study for comparison with conventional CT urography. *Br J Radiol.* 2021 Apr 1;94(1120):20201291. doi: 10.1259/bjr.20201291. Epub 2021 Feb 24. PMID: 33571034; PMCID: PMC8010546.
 21. Shin YJ, Chang W, Ye JC, Kang E, Oh DY, Lee YJ, Park JH, Kim YH. Low-Dose Abdominal CT Using a Deep Learning-Based Denoising Algorithm: A Comparison with CT Reconstructed with Filtered Back Projection or Iterative Reconstruction Algorithm. *Korean J Radiol.* 2020 Mar;21(3):356-364. doi: 10.3348/kjr.2019.0413. PMID: 32090528; PMCID: PMC7039719.
 22. Immonen E, Wong J, Nieminen M, Kekkonen L, Roine S, Törnroos S, Lanca L, Guan F, Metsälä E. The use of deep learning towards dose optimization in low-dose computed

- tomography: A scoping review. *Radiography (Lond)*. 2022 Feb;28(1):208-214. doi: 10.1016/j.radi.2021.07.010. Epub 2021 Jul 27. PMID: 34325998.
23. Greffier, J., Durand, Q., Frandon, J., Si-Mohamed, S., Loisy, M., de Oliveira, F., ... & Dabli, D. (2023). Improved image quality and dose reduction in abdominal CT with deep-learning reconstruction algorithm: a phantom study. *European Radiology*, 33(1), 699-710.
24. Greffier J, Durand Q, Serrand C, Sales R, de Oliveira F, Beregi JP, Dabli D, Frandon J. First Results of a New Deep Learning Reconstruction Algorithm on Image Quality and Liver Metastasis Conspicuity for Abdominal Low-Dose CT. *Diagnostics (Basel)*. 2023 Mar 20;13(6):1182. doi: 10.3390/diagnostics13061182. PMID: 36980490; PMCID: PMC10047497.
25. Greffier J, Dabli D, Frandon J, Hamard A, Belaouni A, Akessoul P, Fuamba Y, Le Roy J, Guiu B, Beregi JP. Comparison of two versions of a deep learning image reconstruction algorithm on CT image quality and dose reduction: A phantom study. *Med Phys*. 2021 Oct;48(10):5743-5755. doi: 10.1002/mp.15180. Epub 2021 Sep 1. PMID: 34418110.
26. Greffier J, Frandon J, Durand Q, Kammoun T, Loisy M, Beregi JP, Dabli D. Contribution of an artificial intelligence deep-learning reconstruction algorithm for dose optimization in lumbar spine CT examination: A phantom study. *Diagn Interv Imaging*. 2023 Feb;104(2):76-83. doi: 10.1016/j.diii.2022.08.004. Epub 2022 Sep 11. PMID: 36100524.

Ringraziamenti

Al Prof. Tabarrini, per la disponibilità, la costanza e l'impegno nell'aiutarmi durante quest'ultima tappa del mio percorso universitario.

Alla mia mamma e al mio papà, per avermi sempre spinto ad essere la migliore versione di me stesso, per avermi dato la possibilità di scegliere e di sbagliare con la mia testa senza mai ostacolarmi, facendomi vivere esperienze che porterò con me per tutta la vita.

A mio fratello, che nonostante i tanti anni di differenza non smette mai di fare il tifo per me e di dimostrarmi quanto mi voglia bene, litigheremo spesso ma non vedo l'ora di vedere chi diventerai da grande.

Ai miei nonni, che sono stati come dei secondi genitori, grazie per essere sempre presenti e orgogliosi di me in ogni passo importante della mia vita, che mi avete sempre donato tanto senza chiedere mai niente indietro se non affetto.

A Tolo, senza il quale non so se avrei mai sopportato e finito questa università, che mi ha sempre seguito nei piani di studio più assurdi e con cui ho condiviso i momenti più belli e brutti degli ultimi tre anni, sempre con la stessa capacità di farmi ridere e di far sembrare ogni ostacolo superabile (dovevamo davvero fare medicina).

A Guolo, Ale, Andre con cui ho condiviso gran parte di questa avventura, che hanno condiviso con me non solo infiniti pomeriggi di studio disperato, ma anche sabati sera improbabili, cene discutibili e una finale degli europei magica. Aspettando di tornare da Sofia so già che sarà una vacanza indimenticabile.

A Francesco, probabilmente una delle amicizie più longeve e importanti che ho, nonostante non ci si riesca mai a vedere e le chiacchierate siano sempre meno mi dimostri sempre come possa sempre contare su di te. P.S. le settimane di estate ragazzi passate insieme rimarranno per sempre.