

Deep Denoising:

BENCHMARKING DI ARCHITETTURE PROFONDE

Relatori:

Prof. Donato Impedovo
Dott. Davide Veneto

Laureando:

Alessandro Congedo



Dipartimento
di
Informatica

Table of Contents

03	INTRODUZIONE	05	STATO DELL'ARTE	10	DATASET
12	RUMORE	13	TECNOLOGIE	14	METRICHE
15	RISULTATI	23	ANALISI FINALE	25	SVILUPPI FUTURI

Introduzione

Con l'avvento dei nuovi sistemi di riconoscimento e autenticazione biometrica, la necessità di immagini di **alta qualità**, prive di distorsioni, si è fatta sempre più pressante.

In questo contesto, le tecniche di **deep denoising** emergono come strumenti cruciali per migliorare l'efficacia di tali sistemi, garantendo la fedeltà e l'integrità dei dati biometrici essenziali per un riconoscimento accurato.

La presente tesi si propone di esplorare il panorama del deep denoising attraverso un'analisi benchmarking di diverse architetture di deep learning, con l'obiettivo di valutarne l'efficienza e l'efficacia nel rimuovere vari tipi di rumore da immagini di impronte digitali.



Problema



La **presenza di rumore** non solo rende più arduo il processo di identificazione ma può anche aumentare il rischio di **falsi positivi o negativi**, minando l'affidabilità dell'intero sistema biometrico.

Prima di essere utilizzate per l'autenticazione o il riconoscimento, queste immagini necessitano pertanto di un accurato "processo" di **pulitura** che ne ripristini la chiarezza e la precisione.

— Soluzione

Il **denoising** emerge come la soluzione ottimale a questo problema, applicando tecniche avanzate di elaborazione delle immagini per filtrare il rumore e migliorare la qualità dell'immagine delle impronte digitali.

Con l'ausilio delle architetture di **deep learning**, questo processo si avvale di modelli capaci di apprendere le caratteristiche distintive delle impronte, distinguendole dalle anomalie indesiderate.



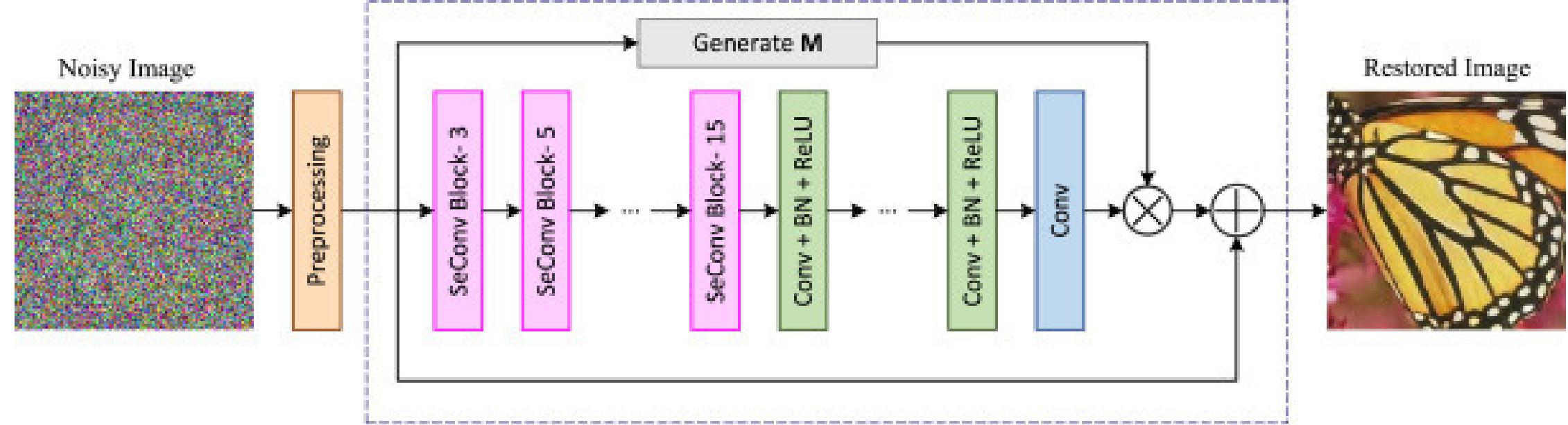
Stato dell'Arte

— Introduzione

Le Convolutional Neural Network (**CNN**), che lavorano con layer di **convoluzione** e di **pooling**, hanno rafforzato le fondamenta gettate dalle metodologie tradizionali, come il Wavelet Thresholding e il Non-Local Means, per l'evoluzione di architetture avanzate di deep learning, ampliando i confini di ciò che è possibile in termini di recupero e miglioramento della qualità delle immagini.

— Architetture Proposte

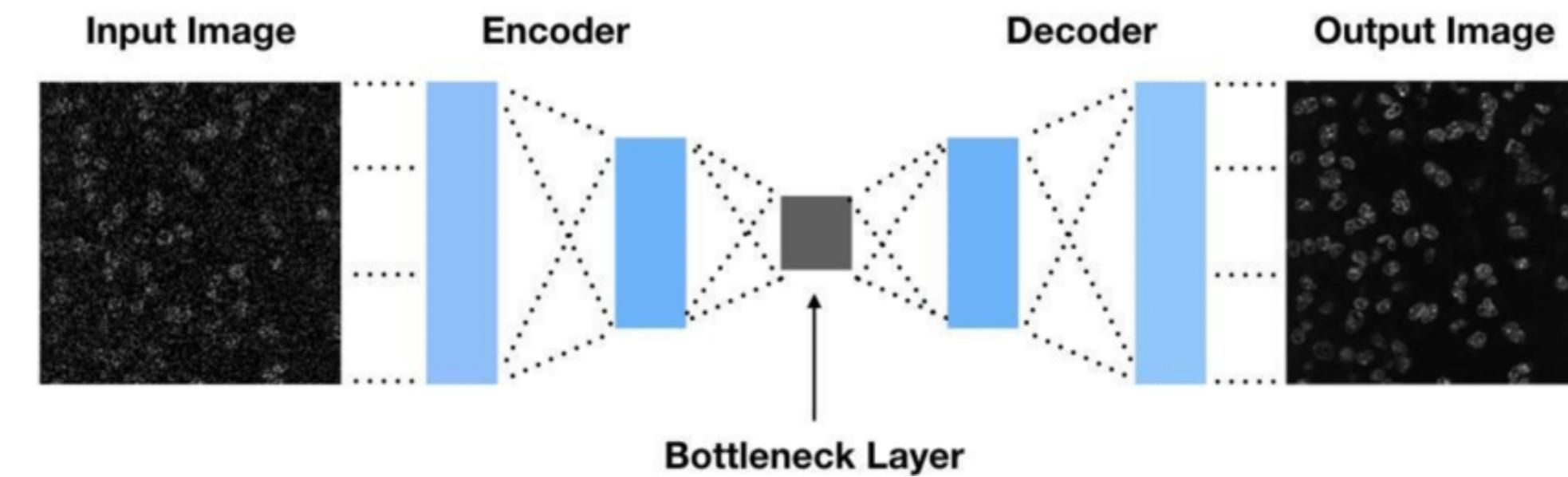
- Selective Convolutional Network (**SeConvNet**)
- Denoising Autoencoder (**DAE**)
- Denoising U-Net (**DUNet**)
- Residual Dense Network (**RDN**)



SeConvNet

Particolarmemente efficace nel trattare immagini multicanale (RGB), **SeConvNet** si distingue per la sua capacità di affrontare efficacemente il rumore salt-and-pepper. Il cuore dell'innovazione di SeConvNet risiede nel suo blocco convoluzionale selettivo (**SeConv Block**), che rappresenta una deviazione strategica dagli approcci convenzionali nel denoising.

Il **SeConv Block** utilizza un approccio selettivo per filtrare i pixel da considerare nel processo di convoluzione. Invece di trattare uniformemente tutti i pixel dell'immagine, il blocco SeConv identifica attivamente i pixel non corrotti dal rumore SAP per partecipare alla convoluzione. Questa selezione si basa sulla mappatura dei pixel rumorosi e non rumorosi dell'immagine in ingresso.



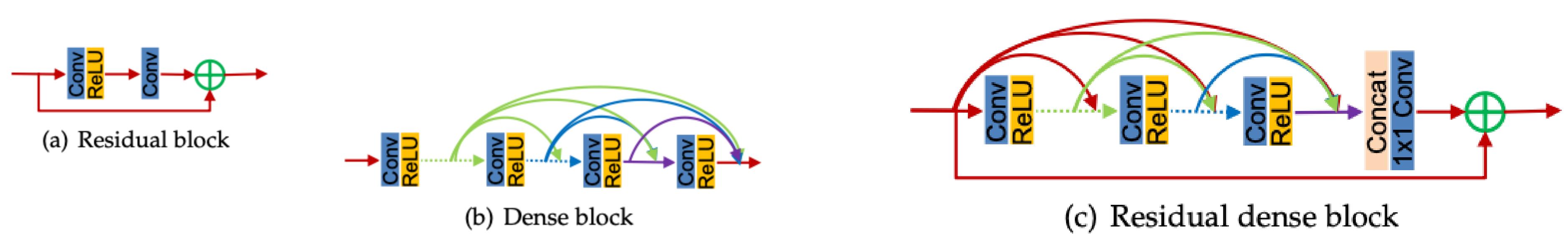
DAE

I **Denoising AutoEncoder** apprendono una rappresentazione compatta (codifica) delle immagini pulite durante il processo di training. Questo processo permette al modello di identificare e "ignorare" il rumore durante la ricostruzione dell'immagine.

Codifica: l'immagine in ingresso (corrotta da rumore) viene compressa in una rappresentazione latente di dimensione ridotta. Questa rappresentazione cattura le caratteristiche essenziali dell'immagine, eliminando idealmente le componenti attribuibili al rumore.

Decodifica: nella fase di decodifica, la rappresentazione latente viene poi espansa per ricostruire l'immagine. L'obiettivo è che l'immagine ricostruita sia una versione denoised dell'originale, avendo "imparato" a omettere il rumore durante la fase di codifica.





RDN

Una RDN incorpora blocchi densamente connessi e connessioni residuali:

Connessione Densa: questa caratteristica è centrale nelle **DenseNet** (Densely Connected Convolutional Networks), dove ogni strato riceve input da tutti gli strati precedenti all'interno di un blocco denso, creando una connessione complessa tra gli strati.

Connessione Residuale: tecnica introdotta con le Reti Neurali Residue (ResNets). Queste connessioni consentono a un'unità di rete di apprendere la funzione residua relativa all'input, piuttosto che apprendere direttamente la mappatura di output desiderata.

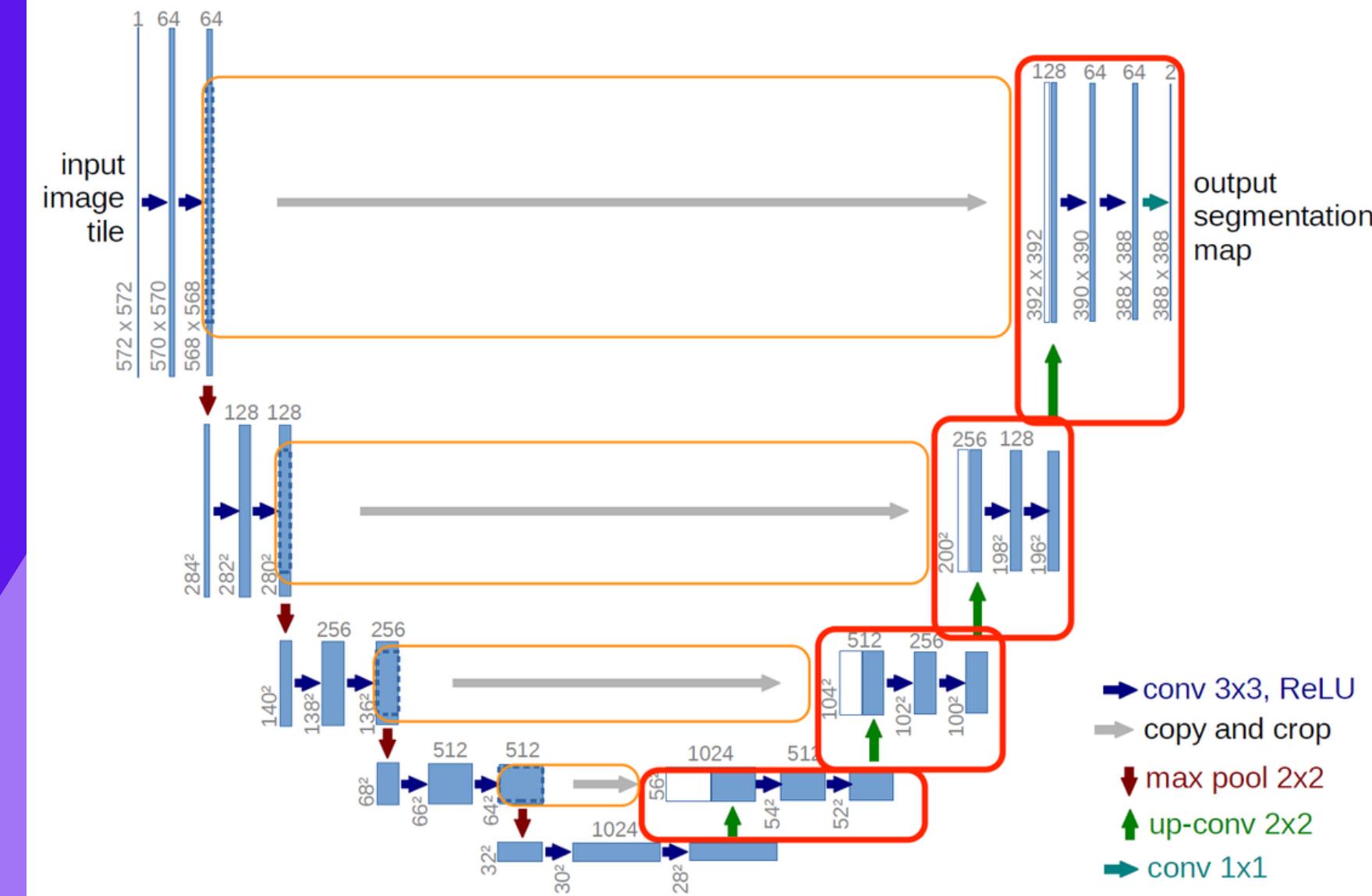
DUNet

La **U-net** è una CNN progettata originariamente per la segmentazione di **immagini biomediche**. La sua architettura encoding-decoding permette una localizzazione precisa.

I layer di convoluzione nel encoder catturano le caratteristiche dell'immagine nelle varie rappresentazioni ridotte delle immagini, ottenute dai layer di pooling.

Le **Skip Connections** permettono il trasferimento diretto di informazioni di contesto ad alta risoluzione dall'encoder al decoder.

Questo aiuta a preservare i dettagli spaziali fini che altrimenti potrebbero essere persi durante il processo di downscaling e upscaling.





Dataset

Il dataset scelto per questo benchmarking è composto da circa 3.000 immagini di impronte digitali in scala di grigio, rese disponibili nel contesto del "**Fingerprint Verification Competition**" dal "**Biometric System Lab**" (Università di Bologna), dal "**Pattern Recognition and Image Processing Laboratory**" (Michigan State University) e dal "**Biometric Test Center**" (San Jose State University).

Il dataset è suddiviso in quattro subset, ognuno dei quali rappresenta una procedura di acquisizione dell'impronta differente, il che è ideale per una corretta generalizzazione dei modelli profondi, e successivamente unito sotto un unico set.

DB1



DB2



DB3

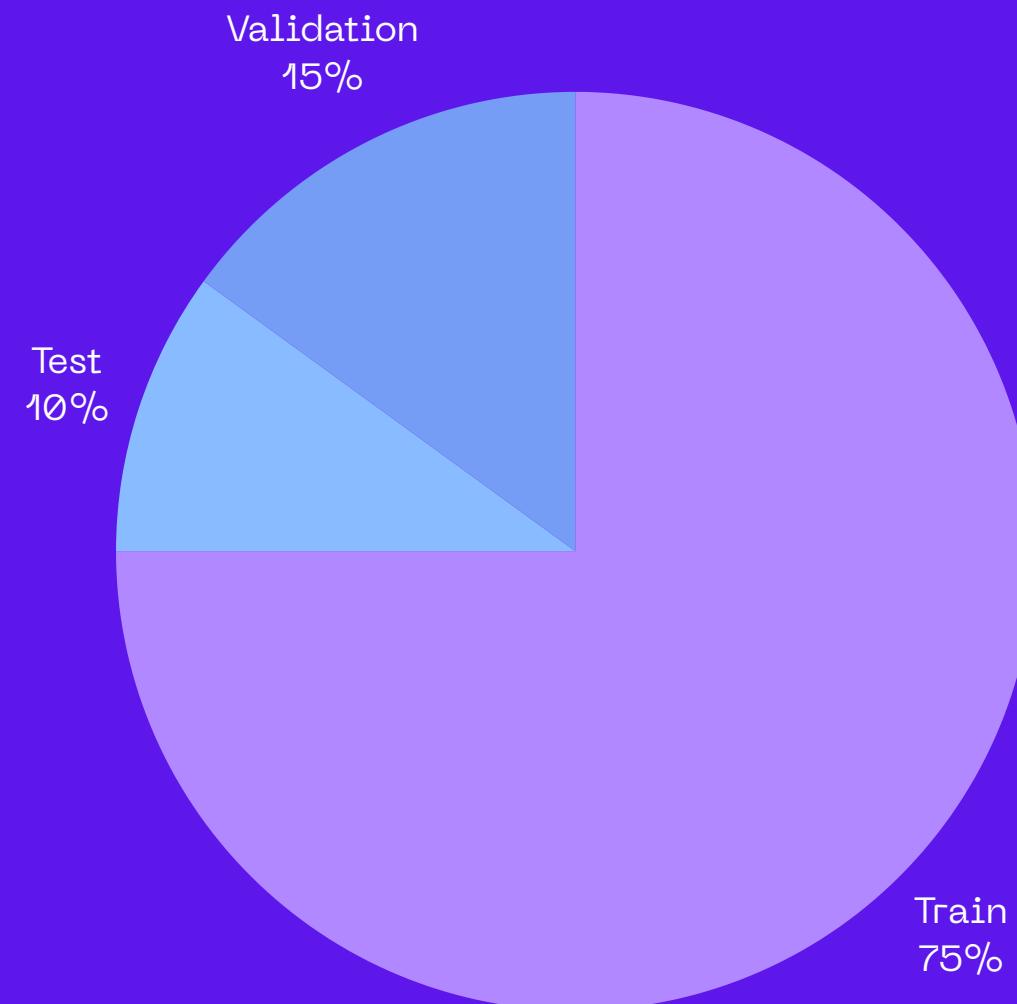


DB4





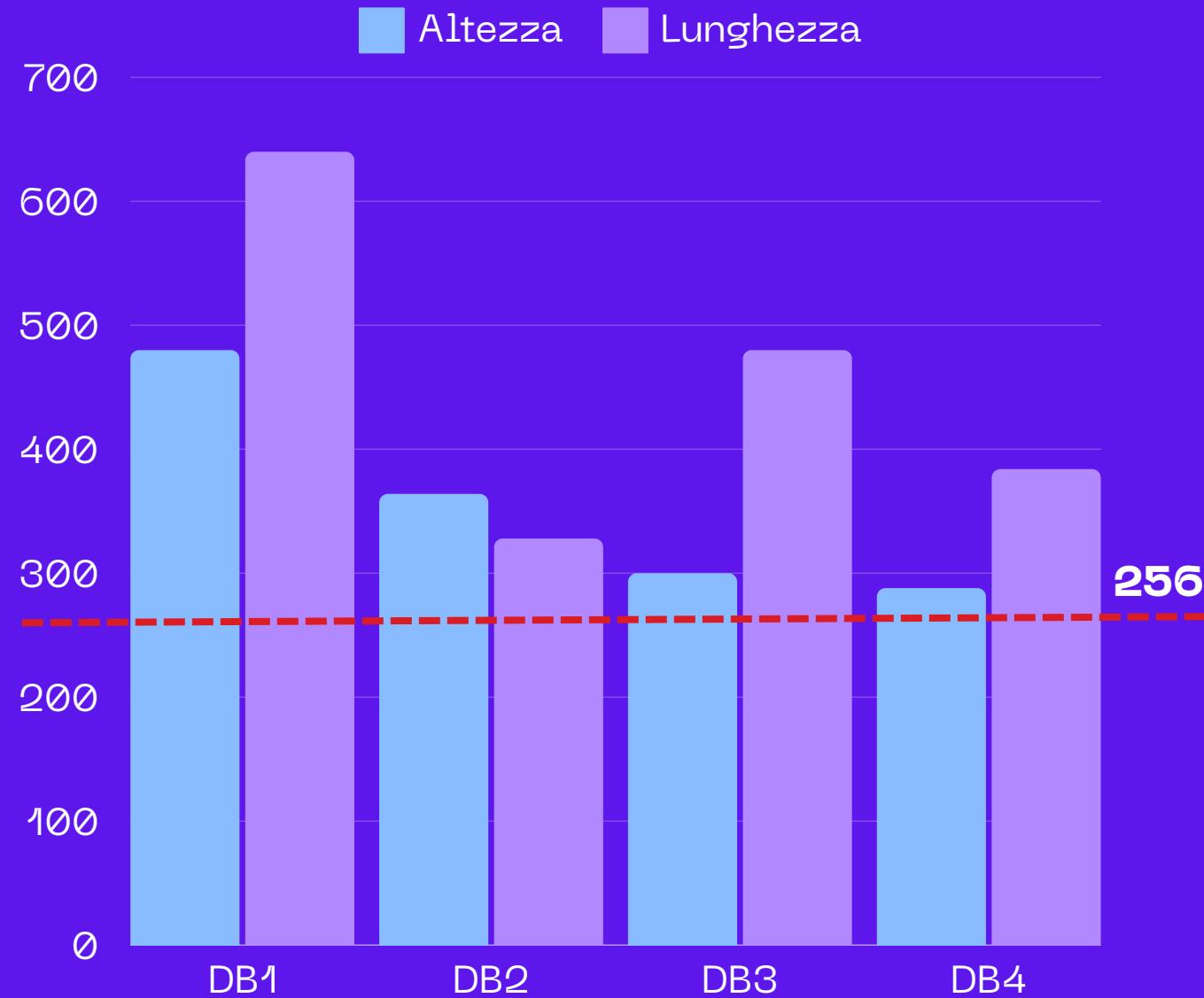
Split & Crop



Il dataset è stato diviso (**split**) in Train set, Test set e Validation set, in linea con le moderne tecniche di data preparation.

Avendo dimensioni diverse, le immagini hanno subito un ritaglio (**crop**) delle dimensioni per uniformare l'input.

Essendo tutte le immagini superiori a **256 pixel**, sia in altezza che in larghezza, è stata utilizzata questa dimensione per ritagliare le immagini (sia in altezza che in lunghezza).





Rumore

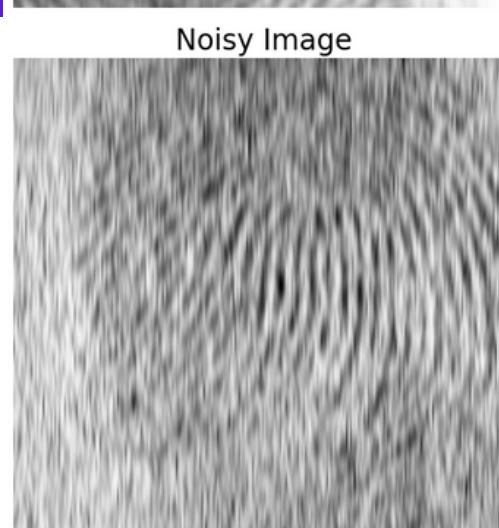
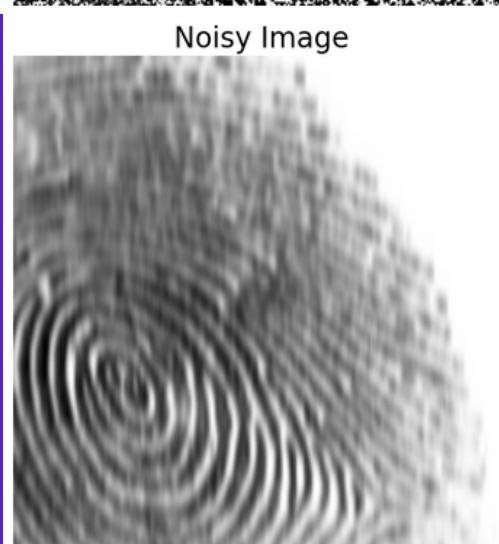
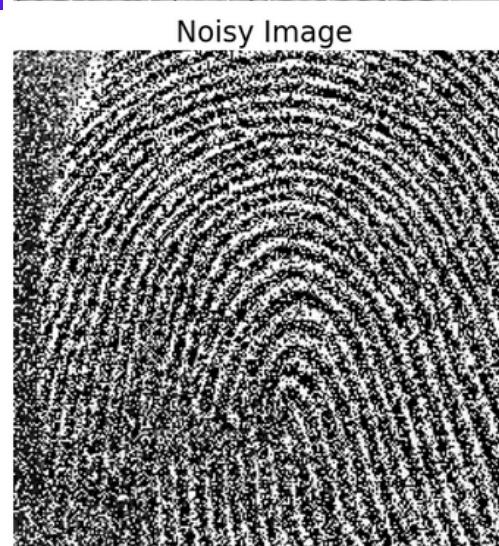
Per analizzare le vere potenzialità di denoising delle architetture scelte, sono stati applicati 4 tipi di rumore direttamente su 4 copie del dataset completo:

- **Salt and Pepper (5%)**
- **Salt and Pepper (50%)**
- **Blur**
- **Salt and Pepper (20%) + Blur**

SAP
+
BLUR

BLUR

SAP
5%
50%



Tecnologie



L'accoppiata **Python + Google Colab** offre un ambiente potente e accessibile per l'addestramento di reti neurali, consentendo agli sviluppatori di sfruttare librerie avanzate e risorse computazionali virtuali direttamente dal browser.



Keras, fungendo da interfaccia di alto livello per **TensorFlow**, semplifica lo sviluppo e l'addestramento di reti neurali con la sua API intuitiva e modulare.

— MSE $\in [0, +\infty)$

La **Mean Squared Error** misura la media dei quadrati delle differenze tra i valori dei pixel delle immagini originali e ricostruite, offrendo una stima quantitativa degli errori.

Un valore di MSE pari a zero indica una perfetta corrispondenza tra le due immagini, mentre valori maggiori indicano una discrepanza.

— PSNR $\in [0, +\infty)$

Basato su MSE, il **Peak Signal Noise Ratio** esprime il rapporto tra l'intensità massima di un segnale e l'impatto del rumore che affligge la sua rappresentazione fedele. Valori più alti di PSNR indicano una maggiore qualità dell'immagine. I valori superiori a 20db sono generalmente considerati accettabili.

— SSIM $\in [-1, 1]$

La **Structure Similarity Index Measure** valuta la percezione della qualità delle immagini basandosi su calcoli di luminosità, contrasto e struttura, offrendo una misura vicina alla percezione umana.

Un valore di SSIM pari a 1 indica una perfetta corrispondenza, mentre valori più vicini a -1 indicano una bassa somiglianza strutturale.

Metriche

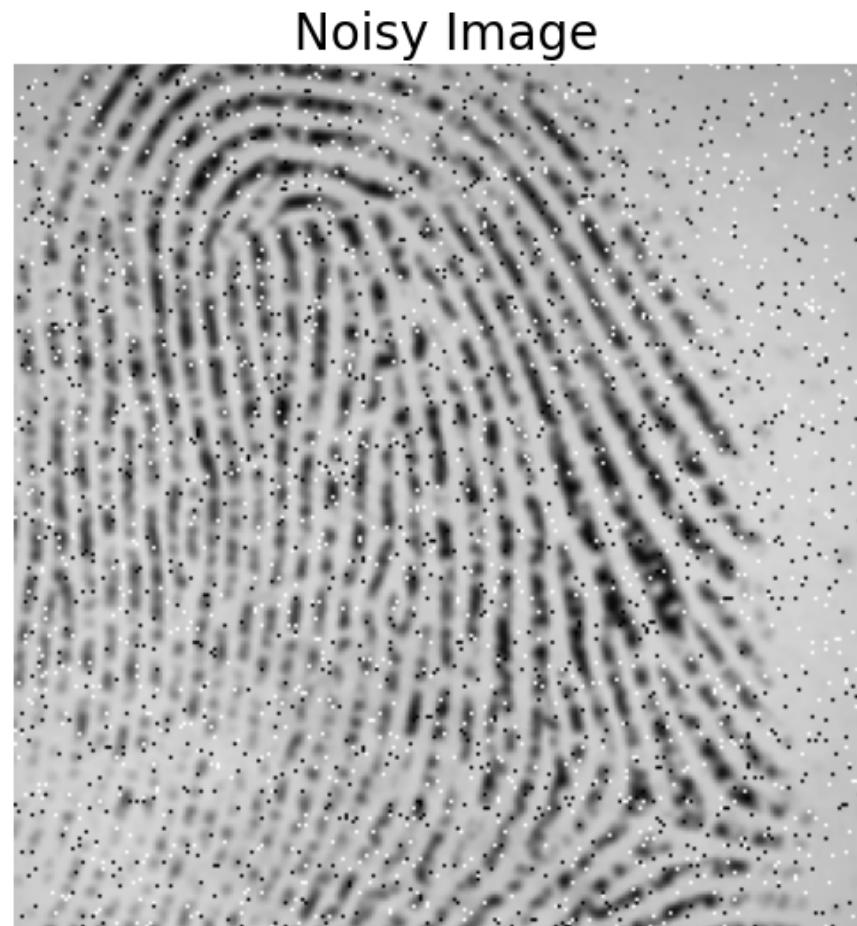
Risultati: SAP 5%

	SeConvNet	DAE	RDN	UNet
MSE	5.7e-03	5.4813e-04	2.5678e-04	3.0195e-04
PSNR	23.5518	36.0727	42.1236	40.4362
SSIM	0.8683	0.9876	0.9963	0.9953

Con un ottimo risultato da tutte le architetture, il rumore "**Salt And Pepper**" al **5%** di intensità rappresenta lo scenario più semplice fra quelli proposti



Risultati: SAP 5%



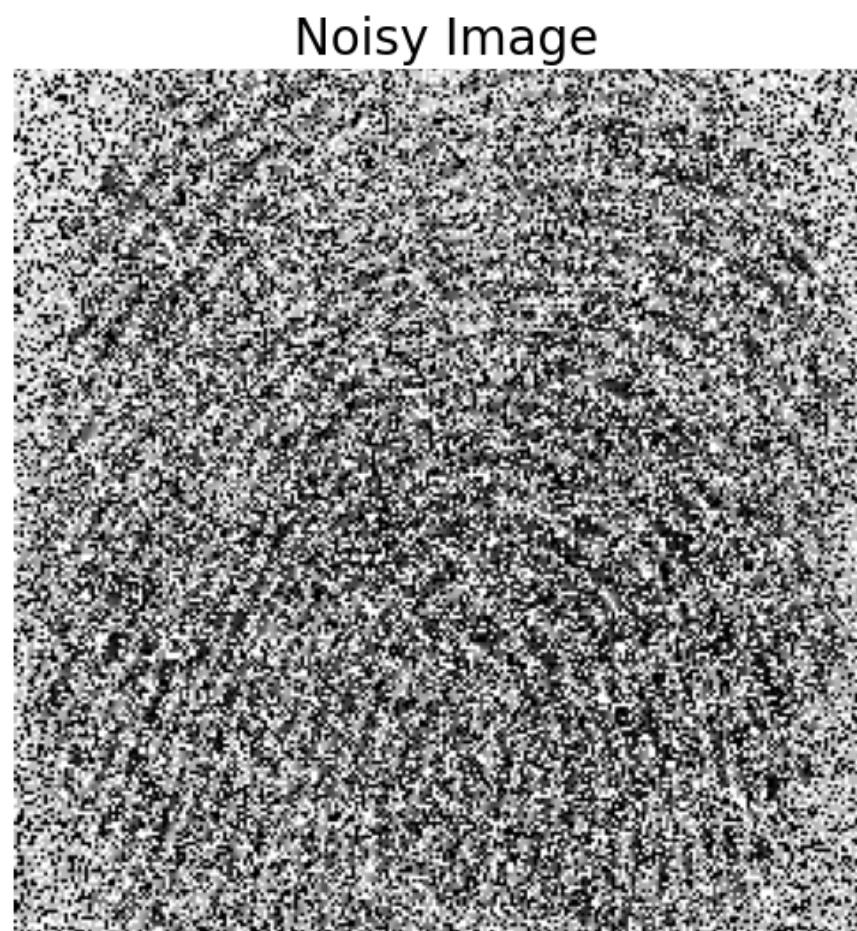
Risultati: SAP 50%

	SeConvNet	DAE	RDN	UNet
MSE	7.065e-02	4.319e-03	5.346e-03	4.269e-03
PSNR	12.5807	27.3627	25.8811	27.5152
SSIM	0.4199	0.9266	0.8962	0.9304

Nel contesto “**Salt And Pepper**” al **50%** di intensità, sia **U-Net** che **DAE** hanno raggiunto degli ottimi risultati, dimostrando grandi capacità di ricostruzione dell’immagine deteriorata.



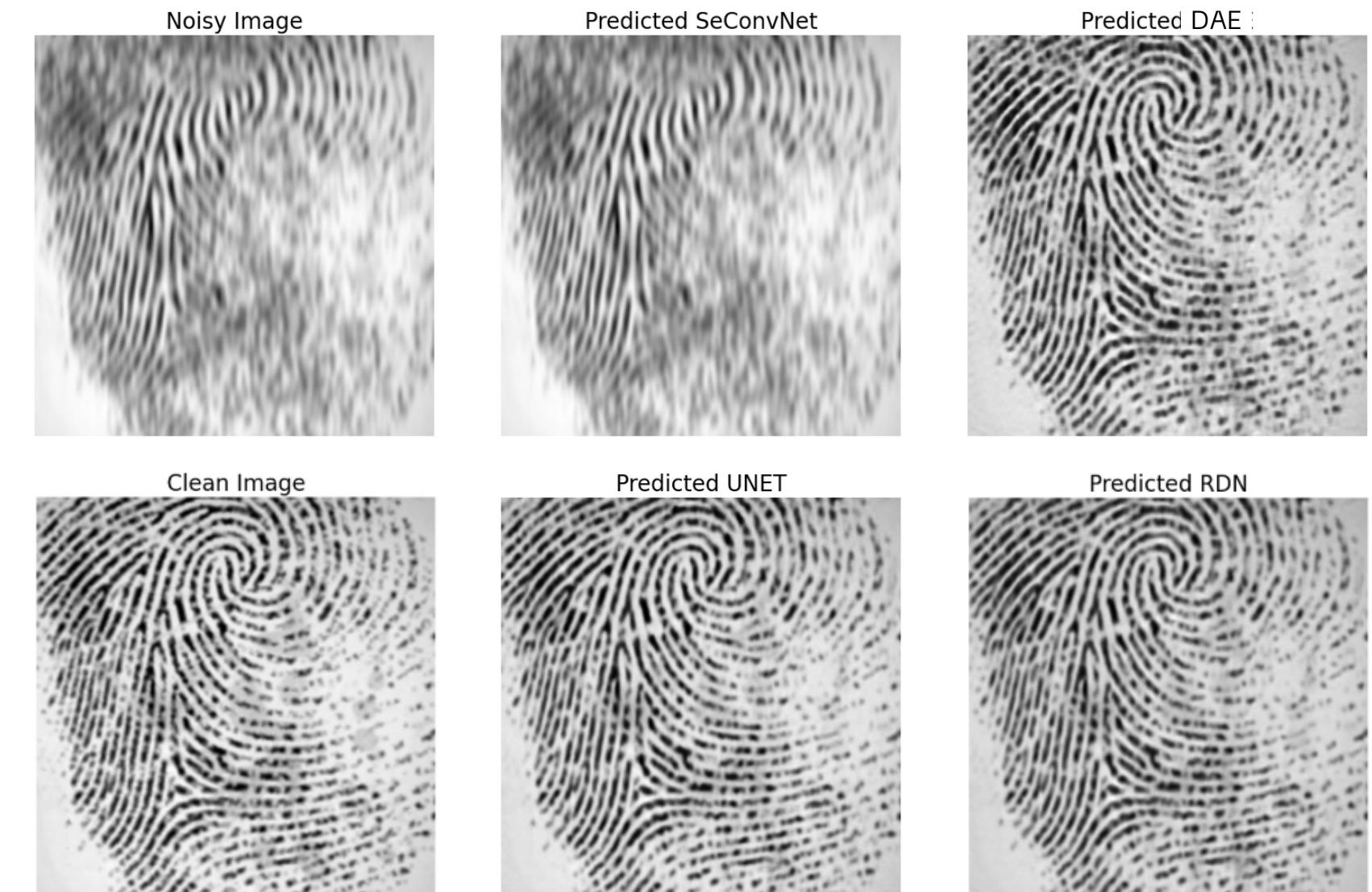
Risultati: SAP 50%



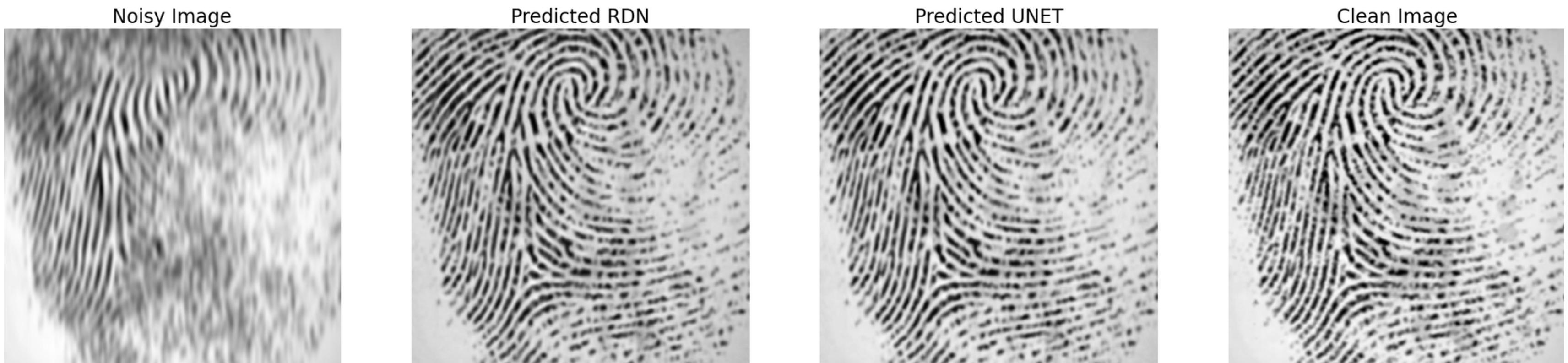
Risultati: BLUR

	SeConvNet	DAE	RDN	UNet
MSE	2.881e-2	5.994e-03	4.351e-03	4.284e-03
PSNR	17.5694	24.5174	25.6795	26.3922
SSIM	0.5327	0.8803	0.9067	0.9153

Il **Blur** rappresenta un altro tipo di rumore, completamente diverso rispetto al SAP: l'esperimento mette alla prova l'adattabilità dei modelli ai vari contesti rumorosi.



Risultati: BLUR

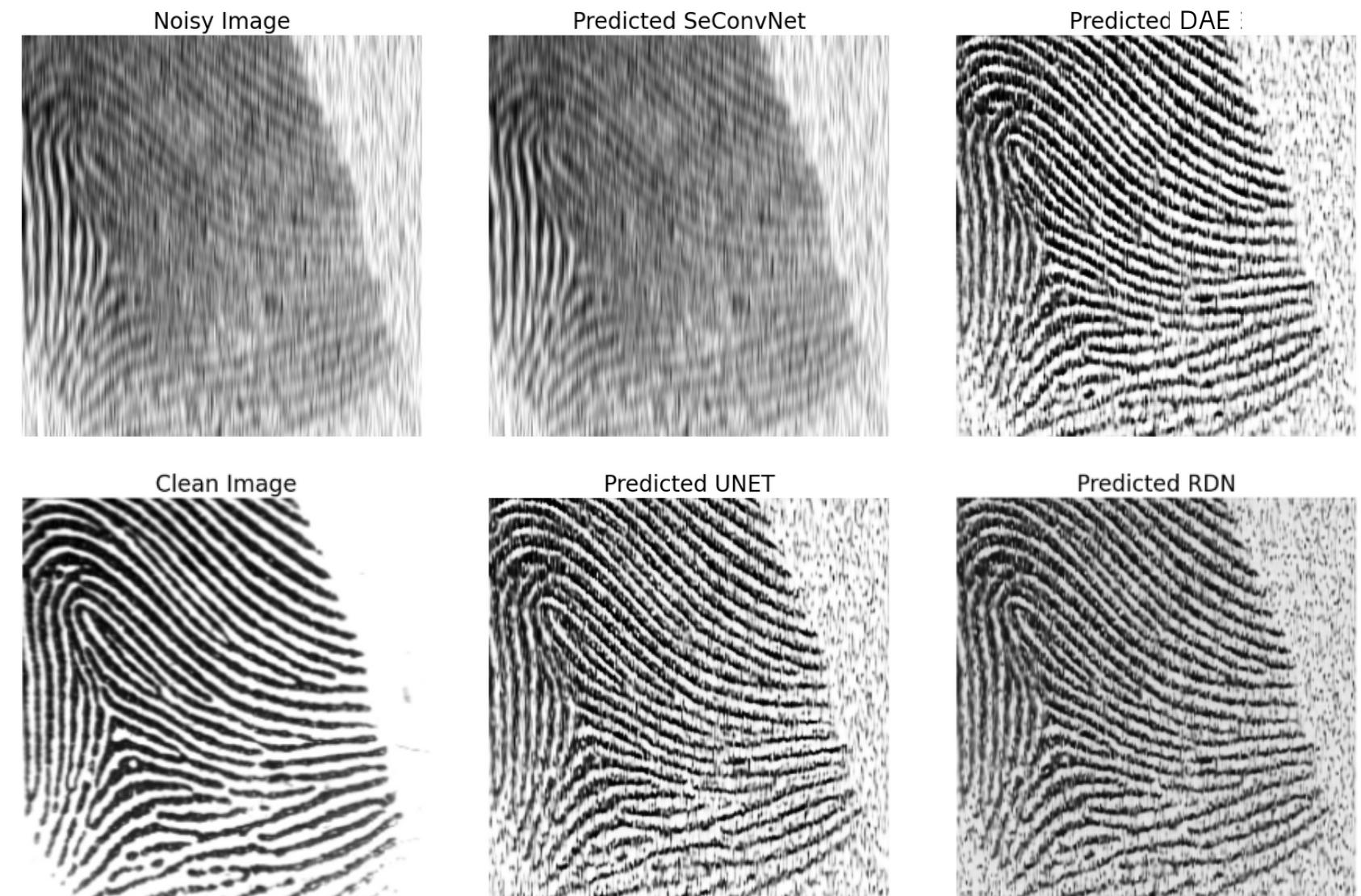


Risultati: **MIXED**

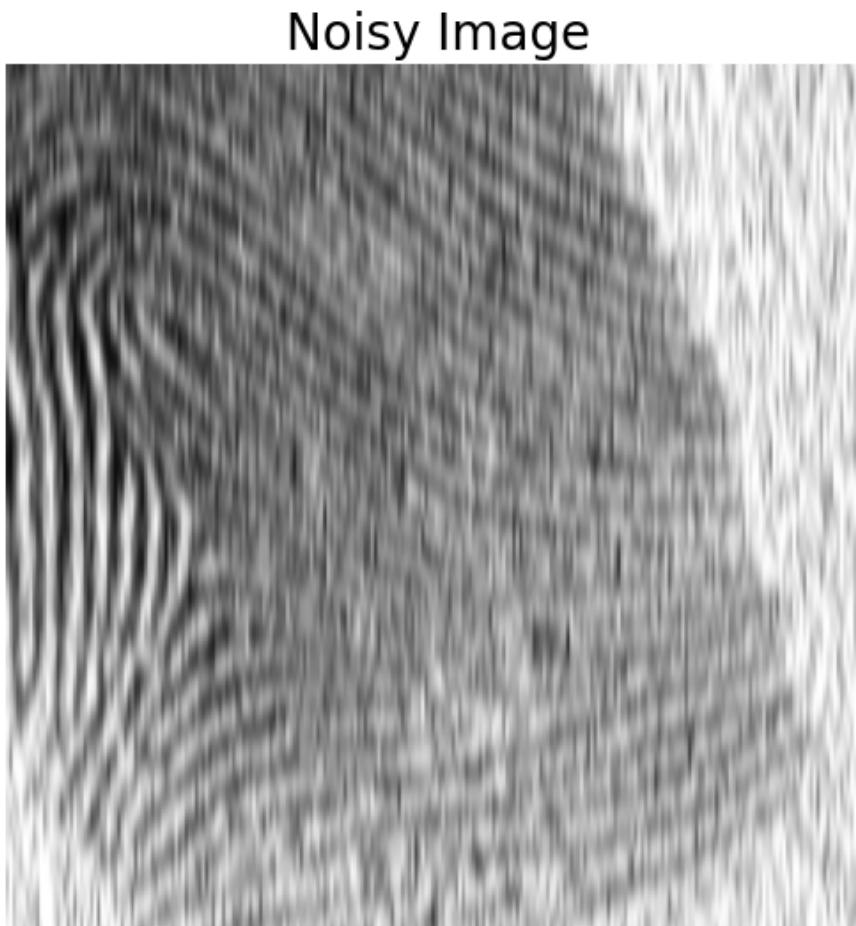
	SeConvNet	DAE	RDN	UNet
MSE	3.794e-02	1.185e-02	1.085e-02	1.016e-02
PSNR	15.5384	20.5199	20.7167	21.2819
SSIM	0.3017	0.7357	0.7499	0.7783

Il rumore "**MIXED (BLUR + SAP 20%)**" ha posto la sfida più complessa per le architetture proposte.

Sebbene le prestazioni abbiano risentito di questa complessità, rivelando un margine di miglioramento, i risultati ottenuti riflettono la potenzialità di evoluzione dei sistemi di denoising.



Risultati: MIXED



Analisi finale



— RDN

La **Residual Dense Network** nonostante abbia riportato ottimi risultati ha richiesto un tempo di addestramento più lungo rispetto alle altre architetture.

— DUnet

L'**U-net** si è dimostrata l'architettura più performante dal punto di vista dei risultati ottenuti e delle risorse impiegate: infatti, è stato il modello che ha impiegato il minor tempo in fase di addestramento.

Analisi finale



— DAE

Nonostante non abbia raggiunto il miglior risultato in nessuna tipologia di rumore, il **Denoising Autoencoder** ha dimostrato grande versatilità e adattabilità nelle varie condizioni proposte, presentando ottimi risultati su ogni fronte.

— SeConvNet

Pur rappresentando un'evoluzione delle canoniche CNN, **SeConvNet** non ha raggiunto i risultati delle architetture concorrenti, performando un'operazione di denoising scadente a confronto.



Sviluppi Futuri

Nel futuro del denoising e del riconoscimento biometrico, prevediamo di concentrarci su diversi aspetti chiave per migliorare ulteriormente la qualità e l'efficacia dei sistemi:

- **Esplorazione di Nuovi Algoritmi:** Testare un ampio spettro di algoritmi di deep learning oltre la **Denoising U-Net** (DUNet) e il **Denoising Autoencoder** (DAE) per identificare metodologie ancora più efficaci nel trattamento del rumore.

- **Validazione con Sistemi Biometrici:** Valutare le prestazioni delle immagini denoised attraverso l'integrazione con sistemi di riconoscimento biometrico esistenti, per testare l'impatto diretto sulla precisione dell'identificazione.
- **Adattamento alle Esigenze Biometriche Specifiche:** Affinare ulteriormente i modelli per accogliere la variabilità dei tratti biometrici umani, incrementando la sicurezza e l'affidabilità dei sistemi di autenticazione.
- **Integrazione con Tecnologie di Acquisizione Avanzate:** Collaborare con le nuove tecnologie di acquisizione delle immagini per sviluppare sistemi biometrici avanzati che mantengano alta qualità dell'immagine anche in condizioni ambientali sfavorevoli.
- **Implementazione su Dispositivi Embedded:** Adattare gli algoritmi di denoising per l'esecuzione efficiente su dispositivi mobili e sistemi embedded, promuovendo l'adozione di soluzioni biometriche pervasive che rispettino la privacy e la protezione dei dati.



UNIVERSITÀ DEGLI STUDI DI BARI "ALDO MORO"

A.A.
2022/2023

Grazie dell'attenzione

Relatori:

Prof. Donato Impedovo
Dott. Davide Veneto

Laureando:

Alessandro Congedo

Dipartimento
di
Informatica

