

Cos'è il noise

Insieme di alterazioni casuali dei valori di intensità dei pixel che non corrispondono all'informazione reale della scena acquisita.

Queste perturbazioni introducono una degradazione della qualità visiva e possono compromettere l'analisi automatica dell'immagine

Cause che creano il noise:

- Rumore del sensore: Imperfezioni dei sensori di acquisizione
- Interferenze elettroniche: Condizioni di illuminazione sfavorevoli
- Limitazioni hardware: Qualità inferiore dei dispositivi di acquisizione (bassa risoluzione)
- Condizioni ambientali: fluttuazioni di temperatura, illuminazione instabile o vibrazioni
- Compressione dei dati: algoritmi di compressione che riducono la dimensione del file ma causano artefatti

Classificazione del rumore

Rumore strutturato → configurazione regolare o pattern nell'immagine

- Rumore sale e pepe → pixel bianchi e neri casuali sparsi nell'immagine
- Rumore a bande (banding noise) → linee orizzontali o verticali regolari nell'immagine

Rumore Non Strutturato → casuale e non segue pattern

- Rumore gaussiano: casuale e segue una distribuzione normale (gaussiana)
- Rumore di Poisson: presente in immagini con scarsa illuminazione, presenta variazione casuali di intensità

Denoising di immagini digitali

Ridurre il rumore senza compromettere alcune caratteristiche essenziali

- Zone uniformi devono restare omogenee
- Bordi devono essere preservati
- Le texture devono essere mantenute
- Non devono essere generati nuovi artefatti

Approcci Principali al Denoising

1. Filtraggio nel dominio spaziale
2. Filtraggio nel dominio delle trasformazioni
3. Metodi avanzati
4. Metodi basati sull'apprendimento automatico

Filtraggio del dominio spaziale

Operazioni sui pixel sfruttando correlazioni locali.

Filtri lineari

Si applicano le operazioni direttamente sui valori dei pixel nell'immagine, sfruttando le correlazioni locali tra un pixel e quelli appartenenti al suo intorno

- Filtro medio: media dei pixel circostanti
- Filtro Gaussiano: ponderazione centrata

Filtri non lineari

- Filtro mediano: valore mediano dei pixel circostanti
- Filtro bilaterale: combina distanza spaziale e intensità per preservare i bordi

Filtraggio nel Dominio delle Trasformazioni

Il filtraggio nel dominio delle trasformazioni non agisce direttamente sui pixel, ma segue tre fasi fondamentali:

1. Trasformazione dell'immagine in un dominio alternativo.
2. Applicazione di filtri per attenuare le componenti associate al rumore.
3. Trasformazione inversa per tornare al dominio spaziale.

Questo approccio consente una separazione più efficace tra informazione utile e rumore.

- Trasformata di Fourier (FT): Rimozione delle alte frequenze.
- Trasformata wavelet (WT): Multi-scala per separare dettagli e rumore.
- Trasformazioni adattive: Si adattano alle caratteristiche specifiche dell'immagine e del rumore

Metodi Avanzati: BM3D

Adotta un approccio iterativo in due stadi costituiti dalle seguenti operazioni di base:

- Block Matching: Regioni simili nell'immagine vengono raggruppate.
- Collaborative Filtering: Applicazione di filtri nel dominio wavelet.
- Aggregation: Combina blocchi ripuliti.

LIMITI dei metodi classici

Nonostante l'elevata efficacia di metodi avanzati come BM3D, gli approcci classici di denoising presentano alcuni limiti strutturali. In particolare, essi si basano su

- ipotesi predefinite sul tipo di rumore
- e su parametri fissati manualmente

risultando poco adattabili a rumori complessi o non stazionari presenti in immagini reali. Inoltre, la loro capacità di preservare dettagli fini e texture complesse è limitata, soprattutto in presenza di livelli elevati di rumore.

Dal punto di vista computazionale, algoritmi come BM3D possono risultare onerosi, rendendo difficile l'applicazione in contesti real-time.

Questi limiti hanno motivato l'adozione di approcci basati su reti neurali convoluzionali, che apprendono direttamente dai dati una rappresentazione più flessibile ed efficace del rumore.

Apprendimento Automatico

Convolutional Neural Networks (CNN): Architettura DnCNN: predice il rumore, lo sottrae per ottenere l'immagine pulita.

- Input: immagine rumorosa
- Output previsto: rumore sintetico aggiunto
- Loss Function: differenza tra il rumore predetto dalla rete e il rumore aggiunto

Generative Adversarial Networks (GAN)

Due reti neurali che lavorano in competizione tra loro:

- Generatore: produce immagini pulite che siano il più simili possibile a quelle reali.
- Discriminatore: valuta l'immagine e cerca di distinguere tra immagini reali e false.

GROUND TRUTH

Definizione di ground truth

La ground truth è il dato di riferimento considerato corretto, che rappresenta l'informazione reale attesa e viene utilizzato per addestrare e valutare le prestazioni di un algoritmo.

Problema dell'assenza della ground truth

In molti contesti reali, incluse le immagini mediche, la ground truth

- non è disponibile
 - o troppo costosa la sua acquisizione
 - o tempi troppi lunghi
 - o rischi del paziente
- o non è acquisibile

Questa assenza rende difficile l'addestramento supervisionato e la valutazione oggettiva dei metodi di denoising.

Introduzione al self-supervised learning e Noise2Void

Per superare la mancanza di ground truth, sono stati introdotti approcci di apprendimento self-supervised, che apprendono direttamente dalle immagini rumorose senza necessità di dati puliti. Un esempio è Noise2Void, che sfrutta la ridondanza spaziale dell'immagine per predire il valore di un pixel a partire dal suo intorno, consentendo il denoising senza immagini di riferimento.

Approcci innovativi

Noise2Noise (N2N):

- Modello allenato su due immagini rumorose della stessa scena.
- Immagini basate sullo stesso contenuto che sono "rumorose" in modi diversi

Vantaggio: Non richiede immagini pulite.

Noise2Void (N2V):

- Lavora su una sola immagine rumorosa.
- Vantaggio: Perfetto per contesti con dati limitati.

CHALLENGE

Abbiamo deciso di condurre un confronto tra 2 dataset per valutare l'efficacia della strategia di denoising basata su U-Net e Noise2Void.

In questo caso permette di analizzare il comportamento del metodo in due scenari differenti:

- uno controllato
con immagini naturali affette da rumore artificiale
- uno realistico
con immagini di microscopia biologica caratterizzate da rumore reale e complesso.

Nel primo caso, la disponibilità del ground truth permette una valutazione quantitativa, verificando la correttezza del denoising.

Nel secondo caso, l'analisi è principalmente qualitativa e orientata alla preservazione delle strutture biologiche, permettendo di valutare l'affidabilità del metodo in un contesto applicativo reale.

- preservazione delle **strutture biologiche**
- assenza di artefatti
- coerenza visiva

JUMP

JUMP Dataset, costituito da immagini di microscopia cellulare ad alta dimensionalità.

Il dataset è ampiamente impiegato in ambito bio-medico ed è caratterizzato da immagini acquisite in condizioni reali, senza una versione pulita di riferimento.

Le immagini del JUMP Dataset presentano rumore reale di acquisizione, dovuto a fattori fisici e strumentali quali

- limitazioni del sensore
- condizioni di illuminazione
- e processi di microscopia.

Questo tipo di rumore è complesso con distribuzioni sintetiche.

Motivazione della scelta del dataset

La scelta del JUMP Dataset è motivata **dall'assenza di ground truth pulita**, condizione tipica delle immagini mediche e biologiche. Ciò rende il dataset particolarmente adatto alla sperimentazione di approcci self-supervised come Noise2Void, che consentono il denoising senza immagini di riferimento.

BSD68

- dataset di 68 immagini naturali pulite
- Il rumore non è acquisito dal sensore, ma aggiunto artificialmente per valutare algoritmi di denoising.

Questo rende l'esperimento controllato

Se un metodo migliora lo notiamo da 2 valori: PSNR e SSIM

PSNR (Peak Signal-to-Noise Ratio) → misura quanto l'immagine denoised è numericamente vicina al ground truth

- Valuta solo pixel per pixel
- No tiene conto di percezione visiva

SSIM → Misura la somiglianza strutturale tra immagine denoised e ground truth

- penalizza la **perdita di dettagli**
- evidenzia se il denoising distrugge bordi e texture

il modello è migliore in uno specifico scenario (rumore controllato) non dice nulla sul comportamento di un rumore reale.

Implementazione del codice

Slide 24

l'U-Net è **implicita** dentro la configurazione **Noise2Void di CAREamics**.

Si definiscono insieme di parametri che specificano l'architettura del modello

Depth

- depth indica il **numero di livelli di downsampling** (encoder)
- di conseguenza anche **2 livelli di upsampling** (decoder)

È una **U-Net poco profonda**, adatta a:

- immagini non troppo complesse
- ridurre overfitting
- dataset piccoli

Numero di feature maps al primo livello

- È il **numero di feature maps nel primo blocco convoluzionale**
- rappresenta la **larghezza iniziale** della rete

Evoluzione dei canali

Con una U-Net standard:

- Livello 1 → **32** feature maps
- Livello 2 → **64** feature maps
- Bottleneck → **128** feature maps

(e poi simmetricamente nel decoder)

👉 Più feature maps = maggiore capacità di rappresentazione, ma più parametri.

TRAINING

Slide 25

- Parte il **training loop**
- Il modello:
 - legge i dati di training (train_source)
 - legge i dati di validazione (val_source)
 - esegue:
 - forward

- calcolo della loss
 - backpropagation
 - aggiornamento dei pesi
- per il numero di epoche definito **nella training configuration**

Predictions

Al termine dell'addestramento, il modello viene utilizzato per generare predizioni sulle immagini rumorose presenti nel set di **testing**

Il metodo `.predict` sfrutta l'**ultimo checkpoint salvato** durante l'addestramento per produrre le immagini denoised.

Nel caso del **Dataset BSD68**, il modello genera predizioni per tutte le 68 immagini del dataset.

Questa fase consente di valutare il comportamento del modello su dati non visti durante l'addestramento e di confrontare le immagini predette con le corrispondenti **Ground Truth**, fornendo una misura oggettiva delle prestazioni del denoising.

Conclusioni → scritte sulla slide