

# Articolo

---

## Breve intro

La colorazione istologica è lo standard per l'esame dei tessuti nella patologia clinica e nella ricerca sulle scienze della vita, che visualizza le strutture tissutali e cellulari utilizzando **coloranti cromatici** o **etichette fluorescenti** per facilitare la valutazione microscopica dei tessuti. Tuttavia, l'attuale flusso di lavoro della colorazione istologica richiede noiose fasi di preparazione dei campioni, infrastrutture di laboratorio specializzate e istotecnologi qualificati, il che lo rende costoso, dispendioso in termini di tempo e non accessibile in contesti con risorse limitate.

Le tecniche di deep learning hanno creato nuove opportunità per rivoluzionare il metodo di colorazione generando digitalmente colorazioni istologiche denominate **Colorazione Virtuale**. Definiamo due grandi categorie di processo:

- Da un'immagine senza etichetta (campione non colorato, label free) → Simulazione di colore
- Da un'immagine di campione già colorato (con un etichetta, labeled) → Altra tipo di colorazione

Alcune delle colorazioni utilizzate e nominate nel paper sono:

- Ematossilina ed Eosina (H&E): crea un contrasto tra i nuclei e la matrice tissutale extracellulare (la ematossilina colora il nucleo in blu/viola e la eosina il citoplasma e le proteine in rosa/arancione)
- Periodic Acid-Schiff (PAS): evidenzia zuccheri e glicoproteine in rosa/magenta
- Tricromica di Masson: evidenzia in blu/verde il collagene, in rosso il citoplasma e i muscoli e in nero/marrone i nuclei

Alcune delle soluzioni e metodologie applicate per ovviare ad alcune problematiche della colorazione chimica sono:

- **Imaging autofluorescente**: sfrutta la **fluorescenza intrinseca** di alcune biomolecole presenti nei tessuti biologici quando vengono eccitate da una sorgente luminosa specifica (tipicamente ultravioletta o blu). E' utile per identificare alterazioni metaboliche senza colorazioni.
- **Imaging di fase quantitativa (QPI)**: permette di visualizzare e quantificare le variazioni dell'indice di rifrazione all'interno di un campione biologico. È particolarmente utile per analizzare cellule vive senza colorazioni. La luce, attraversando un campione biologico trasparente (come cellule o tessuti), subisce variazioni di fase dovute a differenze nell'indice di rifrazione. Il QPI misura queste variazioni di fase e le converte in immagini ad alto contrasto, che rappresentano dettagli strutturali delle cellule e dei tessuti. Permette di ottenere misure quantitative della struttura cellulare, come lo spessore cellulare, l'indice di rifrazione locale e il volume e densità ottica.

---

## Sviluppo del modello

**La parte "Colorazione virtuale senza etichetta" sembra importante e con diversi riferimenti**

Nelle impostazioni di **addestramento supervisionato**, sono necessarie coppie **di immagini di input e ground truth perfettamente incrociate** per addestrare una rete di colorazione virtuale delle immagini; pertanto sono solitamente richiesti modelli di registrazione delle immagini multi-stadio (tecniche per allineare

meglio le immagini di input e ground truth prima dell'addestramento, allineare nel senso che l'immagine di input e di ground truth siano perfettamente sovrapponibili pixel per pixel) o di generazione di dati pre-addestrati (se non ci sono abbastanza coppie reali si possono generare immagini artificiali) per generare immagini di addestramento ben abbinate. D'altro canto, nelle impostazioni di **addestramento non supervisionato**, le **immagini dai domini di input e ground truth non sono necessariamente accoppiate**. Ciò consente di risparmiare sforzi nell'elaborazione preliminare dei dati, tuttavia aumenta la complessità dell'architettura di rete e del programma di addestramento. I framework di apprendimento basati sulla coerenza del ciclo (ad esempio **CycleGAN**) sono comunemente utilizzati in scenari di formazione non supervisionati, che imparano a mappare la distribuzione delle immagini di input nel dominio di verità di base, abbinando il colore e il contrasto.

L'uso del deep learning per ottenere con successo la colorazione virtuale di campioni di tessuto senza etichetta utilizzando immagini di autofluorescenza è stato dimostrato da Rivenson et al, in cui reti neurali profonde sono state addestrate per trasformare le immagini di varie sezioni di tessuto non colorate, ad esempio ghiandole salivari, tiroide, fegato e polmone, in più colorazioni istologiche tra cui H&E, MT e colorazione all'argento di Jones, che corrispondono strettamente alle immagini in campo chiaro delle stesse sezioni di tessuto dopo la colorazione istochimica standard. Negli ultimi anni, sono stati condotti diversi studi per espandere ulteriormente questa tecnica di colorazione virtuale senza etichetta.

---

Più tipi di colorazioni istologiche sono state replicate con successo utilizzando diversi meccanismi di contrasto delle immagini su vari tipi di campioni, il che ha notevolmente arricchito le aree di applicazione dei metodi di colorazione virtuale. Inoltre, aggiungendo matrici di colorazione digitale personalizzate alle immagini di autofluorescenza e utilizzando la loro combinazione come input della rete neurale, Zhang et al. hanno ottenuto colorazioni istologiche microstrutturate e multiplexate sulla stessa sezione di tessuto con una singola rete, il che non è fattibile con il tradizionale flusso di lavoro di colorazione istochimica.

Sintassi: **1. \_"Più tipi di colorazioni istologiche sono state replicate con successo utilizzando diversi meccanismi di contrasto delle immagini su vari tipi di campioni"\_**

- I metodi tradizionali di colorazione istologica (H&E, PAS, Tricromica di Masson, ecc.) sono stati simulati con successo utilizzando tecniche di imaging e modelli di intelligenza artificiale.
- Questo è stato possibile grazie a **diversi meccanismi di contrasto** nelle immagini, come:
  - **Imaging autofluorescente** (senza bisogno di coloranti)
  - **Imaging di fase quantitativa (QPI)** per analizzare variazioni dell'indice di rifrazione nei tessuti
  - **Tecniche ottiche avanzate** per estrarre informazioni strutturali dai campioni.
- Il successo nell'emulare diverse colorazioni ha ampliato il campo di applicazione della colorazione virtuale, rendendola utile in vari settori della diagnostica e della ricerca.

**2. \_"Aggiungendo matrici di colorazione digitale personalizzate alle immagini di autofluorescenza e utilizzando la loro combinazione come input della rete neurale..."\_**

- Per migliorare la colorazione virtuale, sono state **aggiunte matrici digitali personalizzate** che rappresentano diverse colorazioni istologiche.
- Queste matrici digitali vengono applicate alle immagini di autofluorescenza che già contengono informazioni strutturali senza bisogno di coloranti.
- La combinazione di queste informazioni viene utilizzata come input per una rete neurale, che apprende a generare immagini istologiche colorate.

### 3. "... Zhang et al. hanno ottenuto colorazioni istologiche microstrutturate e multiplexate sulla stessa sezione di tessuto con una singola rete" \_

- **Colorazioni microstrutturate:** il modello AI è in grado di generare colorazioni dettagliate a livello microscopico, simulando con precisione le tecniche istochimiche tradizionali.
- **Colorazioni multiplexate:** la stessa sezione di tessuto può essere analizzata con più colorazioni simultaneamente, cosa impossibile con le tecniche tradizionali (che richiedono sezioni separate per ogni colorazione).

### 4. "... il che non è fattibile con il tradizionale flusso di lavoro di colorazione istochimica"

- Nella colorazione istochimica classica, ogni sezione di tessuto può essere trattata con una sola colorazione alla volta.
- Se si vuole applicare più colorazioni (es. H&E + PAS + Masson), servono sezioni di tessuto diverse.
- Con la colorazione virtuale basata su AI, invece, è possibile ottenere più colorazioni sulla stessa sezione di tessuto, migliorando l'efficienza e riducendo il consumo di campioni biologici.

---

Pertanto, oltre alle colorazioni istochimiche standard come H&E e MT, le immagini di autofluorescenza del tessuto privo di etichetta possono essere utilizzate per generare colorazioni molecolari più complesse, ad esempio evidenziando un'espressione proteica specifica, come attualmente fatto dai protocolli di colorazione IHC convenzionali comunemente impiegati nei laboratori di istologia.

Sebbene potente, la microscopia autofluorescente non è l'unica modalità di imaging che consente la colorazione virtuale senza etichetta. Sono state esplorate diverse modalità di imaging che forniscono contrasto per campioni biologici non etichettati per la colorazione virtuale. Ad esempio, il QPI è stato utilizzato anche per la colorazione virtuale. Rivenson et al. hanno utilizzato le immagini di fase quantitative di varie sezioni di tessuto senza etichetta e le hanno trasformate in colorazioni virtuali H&E, Jones e MT utilizzando reti neurali convoluzionali, abbinando le loro controparti colorate istochimicamente in termini di qualità della colorazione.

---

## Preparazione dei training data

L'addestramento dei modelli di colorazione virtuale sopra menzionati richiede solitamente dati di immagine raccolti sia dal dominio di input che da quello di destinazione (ground truth) in modo che il modello possa essere addestrato per sfruttare e tradurre le informazioni dal dominio di input al dominio di destinazione. Tra la raccolta di dati di immagine grezzi e l'addestramento dei modelli di colorazione virtuale, sono necessarie fasi di pre-elaborazione delle immagini per preparare i set di dati per l'apprendimento corretto della trasformazione delle immagini. Queste fasi di pre-elaborazione dei dati si concentrano principalmente sulla registrazione incrociata delle coppie di immagini di input e di destinazione, che è essenziale per i framework di apprendimento supervisionato e per eliminare valori anomali imprevisti, come coppie di immagini non allineate e artefatti di colorazione.

Utilizzando metodi di normalizzazione dei dati adeguati, eventuali deviazioni tra le immagini potrebbero essere ridotte al minimo, in modo che la distribuzione statistica dei dati dell'immagine sia confinata entro un certo intervallo/dominio per promuovere l'apprendimento delle attività di colorazione virtuale.

La cross-registrazione delle coppie di immagini di input e target è comunemente adottata in framework basati sull'apprendimento supervisionato. Un esempio di tale processo di registrazione è stato riportato nel lavoro di

Rivenson et al. PS: probabilmente parliamo di uno di questi quattro lavori:

- 18 (penso sia questo)
  - 19
  - 20
  - 25 sulla colorazione virtuale di immagini di autofluorescenza, in cui è stato implementato un algoritmo di registrazione di immagini multi-modello.
1. Questo algoritmo inizia con una registrazione grossolana delle immagini di autofluorescenza della sezione di tessuto senza etichetta rispetto alle immagini in campo chiaro delle stesse sezioni di tessuto dopo che è stato completato il corrispondente processo di colorazione istochimica, in cui i FOV approssimativamente abbinati di entrambe le modalità di imaging sono stati estratti cercando il punteggio di cross-correlazione più alto.
  2. Quindi è stata stimata una trasformazione affine abbinando i vettori di caratteristiche (descrittori) tra le immagini estratte colorate istologicamente e le immagini di autofluorescenza, che è stata quindi applicata alle immagini colorate per correggere eventuali cambiamenti di scala o rotazione
  3. Nell'ultimo passaggio di registrazione delle immagini più fine, una rete di colorazione virtuale è stata prima addestrata attraverso un basso numero di iterazioni con le immagini approssimativamente abbinate per apprendere la mappatura dei colori
  4. Quindi il modello pseudo addestrato è stato applicato alle immagini di autofluorescenza per assistere la registrazione delle caratteristiche locali utilizzando un algoritmo di registrazione piramidale elastico, che ha aiutato a raggiungere un'accuratezza di co-registrazione a livello di pixel tra le immagini di autofluorescenza di sezioni di tessuto senza etichetta (immagini di input) e le loro corrispondenti immagini di verità di base colorate istochimicamente

Per le immagini microscopiche senza etichetta, il problema dello spostamento di dominio è solitamente osservato come variazioni di imaging che si verificano in diverse condizioni sperimentali. Ciò potrebbe essere causato, ad esempio, da hardware/impostazioni di imaging diversi, ambienti di acquisizione delle immagini incoerenti e variazioni delle caratteristiche del campione o dei protocolli di preparazione del campione. Per affrontare questo problema di spostamento di dominio, la normalizzazione delle immagini viene spesso applicata alle immagini di input senza etichetta prima di inserirle in una rete neurale di colorazione virtuale. Ad esempio, per evitare variazioni di intensità causate da potenziale fotodecolorazione nell'imaging di autofluorescenza, Rivenson et al. hanno normalizzato le immagini di input di autofluorescenza sottraendo il valore medio sull'intero vetrino di tessuto e dividendolo per la deviazione standard dei valori dei pixel. In alternativa, per attenuare queste variazioni e migliorare il contrasto dell'immagine, alcuni lavori hanno saturato l'1% superiore e l'1% inferiore dei pixel. Pradhan et al. hanno anche riferito che la normalizzazione delle immagini multimodali non lineari (NLM) senza etichetta in un intervallo di valori pixel da -1 a 1 potrebbe evitare grandi moltiplicazioni di numeri durante il processo di addestramento, aiutando con una migliore convergenza di rete.

Oltre a queste variazioni che possono essere mitigate da una corretta normalizzazione, a volte le immagini microscopiche catturate potrebbero essere corrotte con, ad esempio, sfocatura, sfocatura da movimento ed errori di lettura. Ad esempio, Zhang et al. hanno presentato un framework di colorazione virtuale utilizzando immagini di autofluorescenza sfocate come input, in cui una rete di messa a fuoco automatica è stata prima addestrata per mettere a fuoco le immagini sfocate casualmente (non ideali), seguita da una rete di colorazione virtuale (che è addestrata congiuntamente) per generare immagini di tessuto virtualmente

colorate a fuoco; questo è stato utilizzato per accelerare significativamente l'intero imaging della diapositiva poiché in questo caso non è necessaria una messa a fuoco automatica fine durante il processo di scansione del tessuto. Allo stesso modo, anche altre condizioni di imaging non ideali all'estremità di input possono essere mitigate utilizzando reti neurali pre-addestrate.

Problemi di spostamento di dominio esistono anche nelle immagini colorate istologicamente, in genere osservati come colori e contrasti estremamente incoerenti a causa di variazioni di colorazione chimica (da laboratorio a laboratorio o da istotecnologo a istotecnologo). Un metodo comune per eliminare tali variazioni nel set di dati di addestramento è quello di utilizzare algoritmi di separazione delle colorazioni e normalizzazione del colore. Tradizionalmente, questi metodi vengono implementati tramite deconvoluzione del colore e mappatura della densità ottica. Oltre a utilizzare metodi di normalizzazione per unificare il colore e il contrasto delle immagini colorate chimicamente, un'altra direzione per mitigare tali problemi di spostamento di dominio nella verità di base è quella di incorporare queste variazioni nel set di dati di addestramento.

Oltre a questi processi di registrazione e normalizzazione delle immagini, viene comunemente eseguita anche una pulizia dei dati algoritmica o manuale per rimuovere i dati indesiderati che potrebbero fuorviare l'addestramento della rete, come sezioni di tessuto deformate o immagini con contaminanti non tissutali (ad esempio, polvere o bolle d'aria).

---

## Architettura di rete e strategie di formazione

Sono state segnalate varie strutture di rete per la colorazione virtuale, tra cui la rete avversaria generativa (GAN) è uno dei framework più comunemente e ampiamente utilizzati grazie alla sua forte capacità di rappresentazione. Rispetto ai modelli di inferenza non basati su GAN, le GAN possono generare immagini con una risoluzione relativamente più elevata e percettivamente più realistiche.

In un framework GAN, due reti neurali profonde, il Generatore e il Discriminatore, sono ottimizzate in modo simultaneo e competitivo. La rete Generatore impara a eseguire la trasformazione dell'immagine dal dominio di input al dominio di destinazione, che in genere utilizza l'architettura U-Net o le sue varianti. Invece la rete Discriminatore è un classificatore che impara a distinguere tra le immagini colorate virtualmente generate dal Generatore e le immagini target colorate istologicamente. Durante l'addestramento, il Discriminatore esamina le immagini colorate virtualmente e restituisce una perdita avversaria al Generatore, aiutandolo a generare immagini che non possono essere distinte dal Discriminatore. Quando l'addestramento entra in uno stato di equilibrio, il Generatore è in grado di creare immagini colorate virtualmente che non possono essere differenziate dalle immagini colorate istologicamente dal Discriminatore. Tuttavia, nel framework GAN standard in cui il generatore è ottimizzato esclusivamente da una perdita avversaria, il generatore risultante imita solo i colori e i modelli delle immagini target senza apprendere la corrispondenza sottostante tra l'input e le immagini target, con conseguenti gravi allucinazioni su scala microscopica. Per superare questo problema di allucinazione, varie altre funzioni di perdita pixel per pixel, come errore assoluto medio (MAE), errore quadratico medio (MSE), SSIM, perdita di Huber, perdita di Huber invertita e metriche di distanza del colore vengono incorporate nei termini di perdita del generatore (oltre alla perdita del discriminatore) per regolarizzare l'addestramento GAN; questi termini di perdita aggiuntivi vengono calcolati utilizzando le immagini generate virtualmente e la loro corrispondente verità di base (immagini colorate istochimicamente). Inoltre, in alcuni lavori sono stati sfruttati anche termini di regolarizzazione delle immagini come la variazione totale per eliminare o sopprimere diversi tipi di artefatti delle immagini creati dal generatore.

Quando sono disponibili coppie di input e immagini target registrate con precisione, spesso la strategia migliore è addestrare una rete di colorazione virtuale utilizzando uno schema di apprendimento supervisionato, poiché le funzioni di perdita pixel per pixel elencate in precedenza possono essere valutate con precisione per l'ottimizzazione del generatore.

A differenza dell'apprendimento supervisionato, l'addestramento di reti di colorazione virtuale mediante schemi di apprendimento non supervisionato non richiede coppie di immagini incrociate. Uno dei framework di apprendimento non supervisionato più frequentemente utilizzati per la colorazione virtuale è l'architettura CycleGAN e le sue varianti, che consistono in due generatori a cascata addestrati congiuntamente per eseguire le trasformazioni di immagine tra il dominio  $X(x)$  e il dominio  $Y(y)$  in modo ciclico. In un ciclo ciclico, il generatore  $G$  esegue prima la trasformazione dal dominio  $X(x) \rightarrow Y(G(x))$ , seguito dal generatore  $F$  che esegue la trasformazione dal dominio  $Y(G(x)) \rightarrow X(x' = F(G(x)))$ . Analogamente, una trasformazione simmetrica dal dominio  $Y(y) \rightarrow X(F(y))$  e poi al dominio  $Y(y' = G(F(y)))$  viene realizzata in un altro ciclo ciclico. Le perdite di coerenza del ciclo come MAE, MSE, e SSIM sono in genere utilizzate in un tale framework di addestramento CycleGAN per misurare le differenze tra  $x \leftrightarrow x'$  e  $y \leftrightarrow y'$ . Inoltre, viene applicata anche una perdita avversaria su  $x \leftrightarrow F(y)$  e  $y \leftrightarrow G(x)$  per garantire la generazione di immagini realistiche. Quando l'addestramento di CycleGAN converge, il generatore  $G$  è in grado di trasferire le immagini dal dominio  $X$  al dominio  $Y$  mentre il Generatore  $F$  può eseguire inversamente la trasformazione dal dominio  $Y$  al dominio  $X$ . Entrambi questi Generatori possono essere estratti e utilizzati nella fase di inferenza a seconda dell'attività di colorazione virtuale desiderata. Un problema notevole dell'utilizzo di CycleGAN per eseguire attività di colorazione virtuale è la mancata corrispondenza dell'intensità; ad esempio, le immagini di input senza etichetta solitamente hanno uno sfondo scuro al contrario delle immagini colorate istologicamente in campo chiaro con sfondo bianco, il che può causare una sfida per le trasformazioni delle immagini a causa della mancanza di supervisione a livello di pixel. Per superare questo problema, oltre a invertire le intensità delle immagini di input senza etichetta, sono stati adottati altri termini di perdita come la perdita del vincolo di salienza e la perdita della misura dell'indice di similarità strutturale multiscala (MS-SSIM). Sebbene le prestazioni dell'apprendimento non supervisionato siano in generale inferiori a quelle dell'apprendimento supervisionato, esso fornisce comunque una soluzione preziosa nei casi in cui i set di dati di immagini accoppiate non sono accessibili/disponibili per l'addestramento.

---

## Valutazione del modello di colorazione virtuale

Dopo l'addestramento di un modello di colorazione virtuale, la sua validità deve essere valutata attentamente con analisi qualitative e quantitative. Un metodo di valutazione semplice e diretto consiste nel misurare direttamente il grado di concordanza tra le immagini istologiche generate virtualmente e le loro controparti colorate chimicamente (ground truth) utilizzando metriche quantitative standard. Quando sono disponibili immagini di input e ground truth accoppiate, vengono comunemente utilizzate metriche di valutazione pixel per pixel, come SSIM, rapporto segnale/rumore di picco (PSNR), MS-SSIM, MSE e MAE. Quando non sono disponibili immagini di input e ground truth accoppiate, possono essere applicate metriche senza riferimento, come la distanza di inizio di Fréchet (FID) e il punteggio di inizio (IS), che valutano le prestazioni di un modello generativo misurando la similarità statistica delle sue immagini di output e target confrontando le caratteristiche di alto livello estratte utilizzando una rete addestrata. Nonostante l'utilizzo di varie metriche di qualità delle immagini e strumenti di analisi delle caratteristiche, i punteggi algoritmici non possono sempre riflettere accuratamente il valore diagnostico delle immagini istologiche virtuali poiché le caratteristiche patologicamente significative sono complesse e non possono sempre essere descritte in modo esplicito tramite semplici regole numeriche. Pertanto, includere patologi certificati negli studi di casi per valutare le

caratteristiche patologiche importanti e prendere decisioni diagnostiche da immagini generate virtualmente è una parte importante del processo di valutazione.