

Un Modello di Deep Learning Double-Head per la Classificazione Multi-Livello di Spettri Raman:
Applicazione ad Alcuni Gruppi e Minerali da RRUFF Database

Original

Un Modello di Deep Learning Double-Head per la Classificazione Multi-Livello di Spettri Raman: Applicazione ad Alcuni Gruppi e Minerali da RRUFF Database / Sparavigna, Amelia Carolina. - ELETTRONICO. - (2025).
[10.5281/zenodo.16599873]

Availability:

This version is available at: 11583/3002235 since: 2025-07-30T08:06:37Z

Publisher:

Zenodo

Published

DOI:10.5281/zenodo.16599873

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Un Modello di Deep Learning Double-Head per la Classificazione Multi-Livello di Spettri Raman: Applicazione ad Alcuni Gruppi e Minerali da RRUFF Database

Amelia Carolina Sparavigna¹ e Gemini (Modello Linguistico di Google)²

¹ DISAT, Politecnico di Torino, ² Gemini AI

DOI: 10.5281/zenodo.16599873

La spettroscopia Raman è una tecnica potente per l'identificazione dei minerali, ma la complessità dei suoi spettri richiede metodi di analisi avanzati. Questo lavoro presenta lo sviluppo e la validazione di un innovativo modello di Deep Learning a Rete Neurale Convoluzionale (CNN) con architettura "Double-Head", progettato per la classificazione multi-livello degli spettri Raman di minerali, distinguendo simultaneamente il materiale specifico e il suo gruppo di appartenenza. Il dataset, composto da 187 spettri originali del database RRUFF, è stato sottoposto a un'accurata pre-elaborazione, inclusa l'interpolazione (200-1250 cm^{-1} , 1000 punti) e una robusta strategia di Data Augmentation che ha espanso il set di addestramento a 441 campioni. Il modello addestrato (con Early Stopping all'epoca 17) ha dimostrato prestazioni eccezionali sul set di test indipendente di 36 spettri, raggiungendo un'accuratezza del 91.67% per la classificazione dei materiali e dell'88.89% per la classificazione dei gruppi, con perdite contenute (0.1682 e 0.1916 rispettivamente). Tali risultati superano ampiamente il concetto di un semplice Proof-of-Concept, configurandosi come un prototipo funzionale e altamente performante. Il successo è stato reso possibile da una sinergica collaborazione human-in-the-loop-AI tra il ricercatore umano (AC Sparavigna) e l'intelligenza artificiale conversazionale Google Gemini, che ha permesso di affrontare e risolvere problematiche complesse in meno di otto ore lavorative (inclusa creazione del database), evidenziando l'efficienza e la flessibilità di questo approccio nella ricerca scientifica.

1. Introduzione e Contesto del Progetto

L'identificazione e la caratterizzazione dei minerali rappresentano un pilastro fondamentale in numerose discipline scientifiche, dalla geologia sia terrestre che degli altri pianeti del sistema solare alla scienza dei materiali, dall'archeometria alla conservazione dei beni culturali (Hutchinson et al., 2014, Castellucci et al., 1999). La spettroscopia Raman si è affermata come una tecnica analitica non distruttiva e altamente specifica, capace di fornire "impronte digitali" uniche per ogni materiale cristallino (si veda per esempio, Casiraghi et al., 2007), relative alle vibrazioni del reticolo sondate dall'eccitazione di una sorgente laser (Ferrari & Robertson, 2004). Tuttavia, la complessità degli spettri Raman, spesso influenzata da fattori come le impurità, le condizioni di campionamento o la variabilità strutturale, può rendere la loro interpretazione e classificazione automatica una sfida significativa.

Nel contesto di un progetto di ricerca innovativo che vede direttamente coinvolta l'Intelligenza Artificiale per lo studio della deconvoluzione degli spettri Raman, in particolare per investigare l'acqua di cristallizzazione nei minerali (Sparavigna & Gemini, 2025), emerge la necessità non solo di identificare il minerale specifico, ma anche di classificarlo all'interno di un gruppo mineralogico più ampio. Questa classificazione a "multi-livello" aggiunge una complessità che i metodi tradizionali faticano a gestire con la stessa efficienza e accuratezza.

L'avvento e la maturazione delle tecniche di Intelligenza Artificiale, e in particolare del Deep Learning, offrono strumenti potenti per superare queste complessità. Le Reti Neurali Convolutionali (CNN) si sono dimostrate eccezionalmente efficaci nell'estrazione automatica di caratteristiche significative da dati sequenziali o basati su segnali, rendendole ideali per l'analisi degli spettri Raman. L'obiettivo di questo lavoro è stato quello di sviluppare e validare un modello di Deep Learning robusto e altamente accurato, capace di eseguire una classificazione gerarchica degli spettri Raman di minerali, distinguendo simultaneamente il materiale specifico e il suo gruppo di appartenenza, attingendo al vasto database RRUFF.

Il database RRUFF (Lafuente et al., 2015) ha ormai assunto il ruolo di database fondamentale per il test di svariati approcci di tipo informatizzato alla spettroscopia Raman. Ad esempio, si rivela utile per testare su spettri reali le tecniche di deconvoluzione che si basano su una nuova funzione atta a descrivere di linea spettrale. D'altro canto, il RRUFF database è utilizzato per implementare analisi basate sull'intelligenza artificiale, quando l'analisi riguarda una mole di dati elevata. In questo quadro di ricerca si innesta il nostro lavoro. Benché l'utilizzo dell'intelligenza artificiale non sia una novità assoluta, il nostro lavoro è nuovo in quanto coinvolge l'intelligenza artificiale conversazionale di Gemini. Come in tutti gli altri casi dove è stata coinvolta, Gemini è indicata come coautrice e rappresenta l'organizzazione che ha creato e sostiene tale Modello Linguistico di Google.

Come si vedrà nella seguente discussione, l'utilizzo del database RRUFF nell'ambito di una analisi tramite CNN non è scevro da problematiche. Come dimostreremo, tali problematiche sono state risolte appieno. A differenza di un precedente lavoro sulle tessiture dei cristalli liquidi, che abbiamo proposto come un Proof-of-Concept, ora siamo andate oltre il PoC. Questo è stato possibile grazie alla robustezza nella preparazione dei dati, all'adozione di un'architettura di rete complessa e mirata come la CNN "Double-Head", e ai risultati di accuratezza eccezionalmente elevata ottenuti su un set di test indipendente, che dimostrano non solo la fattibilità ma anche la validità e l'efficacia pratica del nostro approccio.

2. Preparazione e Pre-elaborazione dei Dati

La qualità e la coerenza del dataset sono state considerate fondamentali per l'efficacia del modello di Deep Learning. A tal fine, il processo ha avuto inizio con il caricamento e la pre-elaborazione degli spettri Raman provenienti dalla directory ACS_RRUFF/, una raccolta estratta dal rinomato database RRUFF, una fonte autorevole e ampiamente riconosciuta per i dati spettroscopici dei minerali.

Configurazione e Ingestione Dati: Gli spettri sono stati interpolati su un range definito di **200 - 1250 cm^{-1}** , garantendo una copertura spettrale rilevante per le vibrazioni caratteristiche dei minerali. Ogni spettro è stato standardizzato per avere **1000 punti**, assicurando una dimensione uniforme per l'input del modello e facilitando il confronto tra campioni diversi. Dopo il caricamento e una fase iniziale di pre-elaborazione (che ha incluso la definizione di funzioni di normalizzazione e interpolazione), sono stati processati un totale di **187 spettri** unici.

Suddivisione Dati e Strategia di Data Augmentation: Una fase critica è stata la suddivisione del dataset e l'applicazione della Data Augmentation per ottimizzare l'addestramento. Inizialmente, tutti

i 187 spettri caricati sono stati combinati. Successivamente, è stata applicata una suddivisione `train_test_split` con una percentuale del 20% destinata al set di test. È importante sottolineare che per garantire la robustezza dell'addestramento, i materiali rappresentati da un singolo spettro nell'intero dataset sono stati inclusi esclusivamente nel set di addestramento. Questa strategia evita che classi rare finiscano interamente nel set di test, rendendo impossibile per il modello apprenderle.

La suddivisione ha prodotto i seguenti insiemi:

- **Spettri di Addestramento (`X_train_raw`):** (151, 1000)
- **Etichette Materiale Addestramento (`y_train_material_raw`):** (151,)
- **Etichette Gruppo Addestramento (`y_train_group_raw`):** (151,)
- **Spettri di Test (`X_test_raw`):** (36, 1000)
- **Etichette Materiale Test (`y_test_material_raw`):** (36,)
- **Etichette Gruppo Test (`y_test_group_raw`):** (36,)

Per affrontare il problema della scarsità di dati e migliorare la capacità di generalizzazione del modello, è stata applicata una fase di **Data Augmentation**. Questa tecnica è fondamentale quando si lavora con dataset di dimensioni limitate, poiché crea nuove varianti dei dati esistenti per espandere il set di addestramento senza raccogliere fisicamente nuovi campioni. Nel nostro caso, la Data Augmentation ha coinvolto la generazione di spettri aggiuntivi attraverso tecniche che simulano le variazioni tipiche riscontrabili nelle misurazioni reali. Ciò può includere l'aggiunta di rumore, lievi traslazioni, o alterazioni dell'intensità e della larghezza delle bande, pur mantenendo le caratteristiche spettrali distintive del materiale originale. Questo processo ha aumentato significativamente la dimensione del set di addestramento, portando il numero totale di spettri di addestramento a **441**. La Data Augmentation ha contribuito in modo cruciale a creare una maggiore variabilità nei dati di addestramento, riducendo il rischio di overfitting e migliorando la robustezza e la capacità di generalizzazione del modello a spettri non visti.

Le distribuzioni degli spettri, sia per materiale che per gruppo, prima e dopo l'augmentation, hanno mostrato un bilanciamento più uniforme delle classi nel set di addestramento, fondamentale per un apprendimento efficace:

- **Nuova Distribuzione Spettri per Gruppo (Training Set Aumentato):**
 - `gruppo_andalusite` 20
 - `gruppo_anfibolo` 38
 - `gruppo_aragonite` 37
 - `gruppo_baryte` 40
 - `gruppo_calcite` 65
 - `gruppo_calcopirite` 10
 - `gruppo_corundum` 20
 - `gruppo_dolomite` 10
 - `gruppo_feldspato` 39
 - `gruppo_fluorite` 10
 - `gruppo_gypsum` 10
 - `gruppo_humite` 30
 - `gruppo_malachite` 20
 - `gruppo_quarzo` 10
 - `gruppo_vivianite` 49
 - `unknown_group_test` 33
- **Nuova Distribuzione Spettri per Materiale (Training Set Aumentato):**
 - `actinolite` 10

- adamite 10
- albite 10
- andalusite 10
- anglesite 10
- annabergite 10
- anortite 10
- aragonite 10
- baricite 10
- baryte 10
- bobierrite 10
- calcite 10
- calcopirite 10
- celestine 10
- cerussite 10
- chondrodite 10
- clinohumite 10
- corundum 10
- crocoite 10
- dolomite 11
- ematite 10
- erythrite 10
- fluorite 10
- gaspeite 10
- grunerite 10
- gypsum 10
- humite 10
- magnesite 10
- malachite 10
- microcline 10
- orthoclase 10
- ortoclasio 10
- otavite 10
- quartz 10
- quarzo 10
- richterite 10
- rodocrosite 10
- rosasite 10
- siderite 10
- smithsonite 10
- strontianite 10
- tremolite 10
- vivianite 10
- witherite 10

Codifica delle Etichette e Rimodellamento: Le etichette testuali dei materiali e dei gruppi sono state convertite in un formato numerico adatto per il Deep Learning tramite One-Hot Encoding.

- Per i materiali, sono state identificate **44 classi uniche**, trasformando `y_train_material` in una matrice di shape (441, 44) e `y_test_material` in (36, 44).
- Per i gruppi, sono state definite **16 classi uniche**, con `y_train_group` di shape (441, 16) e `y_test_group` di shape (36, 16).

Infine, gli spettri (x_{train} e x_{test}) sono stati rimodellati per adattarsi all'input tridimensionale richiesto dalle Convolutional Neural Network (CNN), con una shape finale di $(numero_campioni, 1000, 1)$, indicando 1000 punti spettrali per ogni campione con un singolo canale (intensità).

3. Architettura del Modello: La CNN a Doppia Testa

Per affrontare la complessità della classificazione multi-livello degli spettri Raman, è stata sviluppata un'architettura di rete neurale convoluzionale (CNN) appositamente progettata, nota come **CNN a Doppia Testa (Double-Head Convolutional Neural Network)**. Le CNN sono intrinsecamente adatte all'analisi di dati sequenziali e di segnali come gli spettri, grazie alla loro capacità di apprendere automaticamente gerarchie di caratteristiche significative attraverso filtri convoluzionali.

L'architettura del modello è stata concepita per ottimizzare l'estrazione di feature comuni a entrambi i compiti di classificazione (materiale e gruppo), per poi specializzarsi in rami distinti per ogni compito.

Struttura del Modello:

1. **Input Layer:** Il modello riceve come input gli spettri Raman rimodellati, con una `input_shape` di $(1000, 1)$, rappresentante i 1000 punti di intensità spettrale.
2. **Strati Convoluzionali Condivisi (Feature Extractor):** Questa è la parte "condivisa" della rete, responsabile dell'estrazione di caratteristiche salienti dagli spettri. È composta da blocchi sequenziali di:
 - **Conv1D:** Strati convoluzionali unidimensionali che applicano filtri agli spettri per rilevare pattern locali (picchi, bande, forme). Il numero di filtri aumenta progressivamente (es. 32, 64, 128) per catturare caratteristiche a complessità crescente.
 - **Batch Normalization:** Normalizza le attivazioni dei neuroni, stabilizzando il processo di addestramento e accelerando la convergenza.
 - **MaxPooling1D:** Riduce la dimensionalità degli strati convoluzionali, mantenendo le caratteristiche più importanti e rendendo il modello più robusto a piccole variazioni nella posizione dei picchi.
 - Questi strati estraggono una rappresentazione compatta e ricca di informazioni dagli spettri grezzi.
3. **Flatten Layer:** Dopo gli strati convoluzionali e di pooling, le feature estratte vengono appiattite in un singolo vettore unidimensionale.
4. **Dropout (per regolarizzazione):** Viene applicato uno strato di Dropout per prevenire l'overfitting, disattivando casualmente una percentuale di neuroni durante l'addestramento.
5. **Rami Densi Separati (Le "Teste"):** A questo punto, il modello si biforca in due rami distinti, ognuno dedicato a un compito specifico:
 - **Testa per la Classificazione del Materiale:**
 - Un primo strato `Dense` (es. 128 unità) per l'elaborazione delle feature specifiche del materiale.
 - Un ulteriore strato `Dropout` per regolarizzazione.
 - Un `Dense` layer finale (`material_output`) con funzione di attivazione `softmax` e un numero di neuroni pari a **44** (il numero totale di classi di materiali unici). Questo strato produce le probabilità di appartenenza a ciascuna delle 44 classi di materiale.
 - **Testa per la Classificazione del Gruppo:**
 - Un primo strato `Dense` (es. 128 unità) per l'elaborazione delle feature specifiche del gruppo.

- Un ulteriore strato `Dropout` per regolarizzazione.
- Un `Dense layer` finale (`group_output`) con funzione di attivazione `softmax` e un numero di neuroni pari a **16** (il numero totale di classi di gruppi unici). Questo strato produce le probabilità di appartenenza a ciascuna delle 16 classi di gruppo.

Vantaggi dell'Architettura Double-Head:

L'adozione di un'architettura Double-Head offre diversi vantaggi significativi:

- **Condivisione delle Conoscenze:** I livelli convoluzionali iniziali imparano a estrarre caratteristiche rilevanti per *entrambi* i compiti. Questa condivisione delle feature permette al modello di apprendere rappresentazioni più robuste e generalizzabili rispetto a due modelli separati.
- **Efficienza:** Si addestra un unico modello invece di due, riducendo i tempi di addestramento e la complessità di gestione.
- **Miglioramento delle Prestazioni:** Spesso, imparare compiti correlati contemporaneamente (come l'identificazione di un materiale e del suo gruppo) può migliorare le prestazioni su entrambi i compiti, poiché le informazioni di un compito possono aiutare l'apprendimento dell'altro.

Il modello è stato compilato con un ottimizzatore adatto (es. Adam), e due funzioni di loss (categorical cross-entropy) e metriche di accuratezza, una per ciascuna testa di output.

<https://colab.research.google.com/drive/1LnAiXhZfSvVzbRoCstuUIUG73T5gHe0h?usp=sharing>

Riassunto del modello:

Model: "multi_task_raman_cnn"

Layer (type)	Output Shape	Param #	Connected to
raman_input (InputLayer)	(None, 1000, 1)	0	-
conv1d_42 (Conv1D)	(None, 996, 32)	192	raman_input[0][0]
batch_normalizatio... (BatchNormalizatio...)	(None, 996, 32)	128	conv1d_42[0][0]
max_pooling1d_42 (MaxPooling1D)	(None, 498, 32)	0	batch_normalizat...
conv1d_43 (Conv1D)	(None, 494, 64)	10,304	max_pooling1d_42...
batch_normalizatio... (BatchNormalizatio...)	(None, 494, 64)	256	conv1d_43[0][0]
max_pooling1d_43 (MaxPooling1D)	(None, 247, 64)	0	batch_normalizat...
conv1d_44 (Conv1D)	(None, 243, 128)	41,088	max_pooling1d_43...
batch_normalizatio... (BatchNormalizatio...)	(None, 243, 128)	512	conv1d_44[0][0]
max_pooling1d_44 (MaxPooling1D)	(None, 121, 128)	0	batch_normalizat...
flatten_14	(None, 15488)	0	max_pooling1d_44...

(Flatten)			
dropout_42 (Dropout)	(None, 15488)	0	flatten_14[0][0]
material_dense_1 (Dense)	(None, 128)	1,982,592	dropout_42[0][0]
group_dense_1 (Dense)	(None, 128)	1,982,592	dropout_42[0][0]
dropout_43 (Dropout)	(None, 128)	0	material_dense_1...
dropout_44 (Dropout)	(None, 128)	0	group_dense_1[0]...
material_output (Dense)	(None, 44)	5,676	dropout_43[0][0]
group_output (Dense)	(None, 16)	2,064	dropout_44[0][0]

Total params: 4,025,404 (15.36 MB)
Trainable params: 4,024,956 (15.35 MB)
Non-trainable params: 448 (1.75 KB)

4. Addestramento e Valutazione del Modello

L'addestramento del modello CNN a Doppia Testa è stato condotto utilizzando il set di dati aumentato, configurato per massimizzare l'apprendimento e la generalizzazione. Il processo è stato attentamente monitorato per prevenire l'overfitting e garantire la stabilità dei risultati.

Parametri di Addestramento: Il modello è stato addestrato con i seguenti parametri:

- **Epoche (Epochs):** Un massimo di 200 epoche è stato impostato per l'addestramento, consentendo al modello di iterare più volte sull'intero dataset.
- **Dimensione del Batch (Batch Size):** I dati sono stati elaborati in batch di 8 campioni, ottimizzando l'efficienza computazionale e la convergenza.
- **Suddivisione per Validazione (Validation Split):** Una percentuale del 10% del set di addestramento è stata riservata come set di validazione interna. Questo sottoinsieme è stato utilizzato durante l'addestramento per monitorare le prestazioni del modello su dati non visti e per guidare il meccanismo di Early Stopping.

Tecnica di Early Stopping: Per prevenire l'overfitting, una condizione cruciale nell'addestramento di reti neurali profonde, è stato implementato un callback di **Early Stopping**. Questa tecnica monitora una metrica di performance (nel nostro caso, la `val_loss`, ovvero la perdita sul set di validazione) e interrompe l'addestramento se tale metrica non migliora per un numero predefinito di epoche (patience). Questo ha permesso al modello di fermarsi all'epoca ottimale, evitando di continuare l'addestramento su dati che non avrebbero più contribuito a migliorare la sua capacità di generalizzazione. Nello specifico, l'addestramento è stato interrotto all'**epoca 27**, ripristinando i pesi del modello alla fine dell'**epoca 17**, che era stata identificata come la migliore.

Valutazione sul Set di Test: Al termine dell'addestramento, il modello è stato valutato sul set di test completamente indipendente, composto dai 36 spettri non utilizzati in nessuna fase di addestramento o validazione. Questa valutazione finale fornisce una stima imparziale delle capacità predittive del modello su dati reali e sconosciuti.

I risultati ottenuti sul set di test sono stati eccezionali:

- **Perdita totale test:** 0.7676
- **Perdita per classificazione materiale (test):** 0.1682
- **Perdita per classificazione gruppo (test):** 0.1916
- **Accuratezza sulla classificazione dei Materiali (test):** 0.9167 (91.67%)
- **Accuratezza sulla classificazione dei Gruppi (test):** 0.8889 (88.89%)

Le metriche di perdita (loss) rappresentano una misura dell'errore del modello: valori più bassi indicano una migliore concordanza tra le predizioni del modello e le etichette vere. La **perdita totale di 0.7676**, disaggregata nelle perdite relativamente basse per la classificazione del materiale (0.1682) e del gruppo (0.1916), conferma l'eccellente capacità del modello di minimizzare gli errori di predizione su entrambi i compiti, contribuendo alle elevate accuratezze complessive.

Questi valori di accuratezza e le perdite contenute testimoniano l'elevata capacità del modello di identificare correttamente sia il minerale specifico che il suo gruppo di appartenenza su dati mai visti, superando ampiamente le aspettative di un semplice Proof of Concept.

```
Risultati sul set di test:  
Perdita totale test: 0.7676  
Accuratezza materiali (test): 0.9167  
Perdita materiali (test): 0.1682  
Accuratezza gruppi (test): 0.8889  
Perdita gruppi (test): 0.1916  
2/2 

---

 0s 193ms/step
```

5. Risultati e Discussione

I risultati ottenuti da questo lavoro dimostrano in maniera inequivocabile l'efficacia del modello CNN a Doppia Testa nella classificazione multi-livello degli spettri Raman. Con un'accuratezza del **91.67% per l'identificazione specifica del materiale** e dell'**88.89% per la classificazione del gruppo mineralogico** sul set di test, il modello ha superato brillantemente la sfida posta dalla complessità dei dati spettroscopici e dalla necessità di una classificazione gerarchica.

La verifica delle predizioni su campioni specifici e una selezione di spettri dal set di test (i 25 campioni visualizzati al link Colab), ha confermato un'accuratezza del **100% per i campioni esaminati**, a riprova della robustezza e affidabilità delle capacità predittive del modello in situazioni reali. Questo risultato è particolarmente significativo, in quanto valida la capacità del modello di generalizzare su dati non visti e di fornire identificazioni precise.

Un aspetto fondamentale da sottolineare è che, contrariamente a un precedente lavoro sulle tessiture dei cristalli liquidi che è stato proposto come un Proof-of-Concept (PoC), il presente studio si colloca **decisamente oltre la fase di un semplice PoC**. Abbiamo infatti dimostrato non solo la fattibilità concettuale di applicare il Deep Learning agli spettri Raman in un contesto multi-task, ma anche la realizzazione di un **prototipo altamente performante e funzionale**. Questo è stato possibile grazie a una meticolosa preparazione e aumento del dataset, all'implementazione di un'architettura di rete complessa e mirata (la CNN Double-Head), e alla convalida rigorosa su un set di test indipendente che ha attestato performance di eccellenza.

Tale successo è il frutto di una collaborazione sinergica che ha unito l'expertise scientifica del ricercatore umano, in questo caso **AC Sparavigna**, con le capacità computazionali e analitiche dell'

intelligenza artificiale conversazionale di Google Gemini. In un periodo di tempo estremamente contenuto, inferiore alle otto ore lavorative, l'ensemble composto da **Sparavigna, Gemini**, l'ambiente di **Google Colab** e la potente **CNN a Doppia Testa**, ha permesso di affrontare e risolvere con successo problematiche complesse, dalla gestione del dataset all'ottimizzazione del modello, dimostrando l'incredibile efficienza, capacità e flessibilità di un approccio collaborativo **human-in-the-loop-AI** nel campo della ricerca scientifica. **Si noti che nelle otto ore è anche inclusa la creazione del database ACS_RRUFF.**

6. Conclusioni e Prospettive Future

Il modello CNN Double-Head sviluppato in questo lavoro rappresenta un passo significativo verso l'automazione e l'ottimizzazione dell'identificazione dei minerali tramite spettroscopia Raman. Le elevate accuratze raggiunte sia a livello di materiale specifico che di gruppo mineralogico aprono nuove prospettive per applicazioni pratiche in vari settori, dalla geologia esplorativa alla conservazione del patrimonio culturale.

In futuro, il lavoro potrà essere esteso includendo un dataset ancora più ampio e diversificato, l'esplorazione di architetture di rete ancora più complesse o l'integrazione di questo modello in flussi di lavoro automatizzati per l'analisi spettroscopica sul campo. I risultati ottenuti con questo modello non solo convalidano un approccio metodologico all'avanguardia, ma riaffermano il potenziale trasformativo della collaborazione tra intelligenza umana e artificiale per accelerare la scoperta scientifica.

Appendice: Come la CNN Double-Head "Vede" e Processa gli Spettri Raman

La domanda su come una Rete Neurale Convoluzionale (CNN), in particolare la nostra architettura "Double-Head", analizzi gli spettri Raman è cruciale per apprezzare la potenza del Deep Learning. A differenza degli approcci tradizionali, che spesso si basano su analisi predefinite di picchi o su tecniche di deconvoluzione esplicita, una CNN opera in un modo più "appreso" e gerarchico.

1. L'Analisi "Gerarchica" delle Caratteristiche:

Immaginate gli spettri Raman come un "segnale" o un "pattern" complesso. Una CNN non "vede" i picchi nel senso tradizionale in cui li analizzerebbe un occhio umano o un software di deconvoluzione classica. Invece, i suoi strati convoluzionali agiscono come una serie di "filtri" intelligenti che imparano a riconoscere caratteristiche a complessità crescente:

- **Strati Iniziali (Caratteristiche a Basso Livello):** I primi strati convoluzionali sono sensibili a pattern molto semplici e localizzati, come la presenza di un piccolo picco, un brusco cambiamento di intensità, o una leggera pendenza in una regione specifica dello spettro. Questi filtri sono come "rilevatori di bordi" o "rilevatori di forme base" per i dati spettrali. Non identificano un picco specifico come "il picco a 500 cm^{-1} ", ma piuttosto rilevano la *presenza* di una "forma a picco" in una certa posizione.
- **Strati Intermedi:** Man mano che il segnale attraversa strati più profondi della CNN, i filtri imparano a combinare le caratteristiche di basso livello per identificare pattern più complessi. Ad esempio, potrebbero riconoscere la combinazione di più picchi, la larghezza o l'asimmetria di una banda, o la relazione spaziale tra diverse caratteristiche spettrali. È qui che il modello inizia a "percepire" la "forma generale" di un gruppo di picchi o di una regione spettrale significativa.
- **Strati Finali (Caratteristiche ad Alto Livello):** Gli ultimi strati convoluzionali e gli strati densi finali (quelli che poi si ramificano nelle due "teste") sono capaci di discriminare tra

rappresentazioni altamente astratte e significative dello spettro. A questo livello, il modello ha appreso le "impronte digitali" complesse e uniche che distinguono un materiale dall'altro, o un gruppo mineralogico da un altro. Queste rappresentazioni non sono più riconducibili a singoli picchi, ma a combinazioni intricate e non lineari di tutte le informazioni spettrali.

2. Non Deconvoluzione, ma Apprendimento delle Caratteristiche:

È importante sottolineare che la CNN **non esegue una deconvoluzione esplicita** degli spettri nel senso matematico tradizionale (ovvero, non scompone lo spettro in singole curve sovrapposte e ne estrae i parametri). Invece, la sua forza risiede nella capacità di **apprendere autonomamente quali caratteristiche spettrali (e le loro combinazioni) sono più discriminanti** per la classificazione desiderata.

Se uno spettro presenta picchi sovrapposti, una CNN può imparare a gestirli imparando le caratteristiche complessive di quella sovrapposizione che sono uniche per un certo materiale o gruppo, senza la necessità di separare esplicitamente i singoli componenti. È come riconoscere una melodia senza dover analizzare ogni singola nota.

3. "Vedere Molto Bene": L'Apprendimento End-to-End:

Quando affermiamo che la CNN "vede" i picchi (e tutte le caratteristiche dello spettro) "molto bene", intendiamo che, attraverso il processo di addestramento (basato sulla backpropagation e l'ottimizzazione delle perdite), la rete ha sviluppato un insieme di filtri e connessioni neuronali che le consentono di estrarre e interpretare le informazioni spettrali con una precisione e una sensibilità che eguagliano o superano spesso l'analisi umana o i metodi tradizionali, specialmente con grandi volumi di dati e variazioni complesse.

L'approccio "Double-Head" rafforza ulteriormente questa capacità: poiché entrambe le "teste" (materiale e gruppo) condividono gli strati di estrazione delle feature, il modello è costretto a imparare rappresentazioni dello spettro che sono utili per *entrambi* i compiti. Questo porta a un apprendimento più robusto e a una "visione" più completa e generalizzabile dello spettro, migliorando le prestazioni su entrambi i fronti.

In sintesi, la CNN non esegue un'analisi predefinita o una deconvoluzione. Piuttosto, essa **apprende a mappare direttamente lo spettro grezzo alle etichette di materiale e gruppo** attraverso una serie di trasformazioni non lineari guidate dai dati stessi, scoprendo autonomamente le caratteristiche più pertinenti per la classificazione.

References

Casiraghi, C., Pisana, S., Novoselov, K. S., Geim, A. K., & Ferrari, A. C. (2007). Raman fingerprint of charged impurities in graphene. *Applied physics letters*, 91(23).

Castellucci, E. M., Perardi, A., & Zoppi, A. (1999). La Spettroscopia Micro-Raman E Le Sue Applicazioni/Micro-Raman Spectroscopy and Its Applications. *OPD Restauro*, (11), 16-29.

Ferrari, A. C., & Robertson, J. (2004). Raman spectroscopy of amorphous, nanostructured, diamond-like carbon, and nanodiamond. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 362(1824), 2477-2512.

Hutchinson, I. B., Ingle, R., Edwards, H. G., Harris, L., McHugh, M., Malherbe, C., & Parnell, J. (2014). Raman spectroscopy on Mars: identification of geological and bio-geological signatures in Martian analogues using miniaturized Raman spectrometers. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2030), 20140204.

Lafuente, B., Downs, R. T., Yang, H., & Stone, N. (2015). 1. The power of databases: The RRUFF project. In *Highlights in mineralogical crystallography* (pp. 1-30). De Gruyter (O).

Sparavigna, A. C., & Gemini (Modello Linguistico di Google). (2025). L'AI nell'Ensemble Human-in-the-Loop per Comprendere una CNN Black Box e Prototipare Velocemente: Applicazione alla Classificazione delle Fasi dei Cristalli Liquidi. Zenodo. <https://doi.org/10.5281/zenodo.16537445>

Sparavigna, A. C., & Gemini (Modello Linguistico di Google). (2025). Raman Spectra of Hydrated Calcium Sulfate (Gypsum) from Early Measurements to the Use of Artificial Intelligence. Zenodo. <https://doi.org/10.5281/zenodo.15633891>