

Oltre la Scatola Nera: L'Emergenza dello Pseudo-Spettro come Archetipo dell'Intelligenza Artificiale per l'Analisi Spettrale Non Supervisionata Dalla Mineralogia all'Astrofisica

Original

Oltre la Scatola Nera: L'Emergenza dello Pseudo-Spettro come Archetipo dell'Intelligenza Artificiale per l'Analisi Spettrale Non Supervisionata Dalla Mineralogia all'Astrofisica / Sparavigna, Amelia Carolina. - ELETTRONICO. - (2026). [10.5281/zenodo.18139563]

Availability:

This version is available at: 11583/3006267 since: 2026-01-03T12:25:09Z

Publisher:

Published

DOI:10.5281/zenodo.18139563

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)

Oltre la Scatola Nera: L'Emergenza dello Pseudo-Spettro come Archetipo dell'Intelligenza Artificiale per l'Analisi Spettrale Non Supervisionata Dalla Mineralogia all'Astrofisica

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

¹ DISAT, Politecnico di Torino, ² Gemini AI

DOI: [10.5281/zenodo.18139563](https://doi.org/10.5281/zenodo.18139563)

Il presente lavoro formalizza il concetto di "**Pseudo-Spettro**", una metodologia innovativa emersa nel corso di una serie di studi condotti nel 2025 sull'applicazione di architetture di Deep Learning (Autoencoder Convoluzionali, Densi e Transformer) alla spettroscopia. Tradizionalmente, gli autoencoder vengono utilizzati per la riduzione della dimensionalità o il denoising; qui, proponiamo un cambio di paradigma: l'uso dello spazio latente per generare firme spettrali archetipiche. Lo pseudo-spettro viene definito come la ricostruzione operata dal decoder a partire dal centroide lineare di un cluster nello spazio latente. Attraverso l'analisi di dataset eterogenei — che spaziano dalla mineralogia (Granati, Feldspati) alla scienza dei materiali (Biochar) fino all'astrofisica (Libreria MILES) — dimostriamo che lo pseudo-spettro non è una semplice media statistica, ma l'interpretazione ideale e priva di rumore che il modello ha appreso di una specifica classe di dati. I risultati evidenziano come questa tecnica permetta di "aprire la scatola nera" dell'intelligenza artificiale, fornendo ai ricercatori uno strumento tangibile per visualizzare la logica decisionale del modello e trasformando l'AI da un mero classificatore a una lente interpretativa per la scoperta scientifica.

Basandoci sulla nostra collaborazione e sui risultati ottenuti nei vari articoli (dai minerali alle stelle, fino al biochar), possiamo formalizzare in modo rigoroso il concetto di **Pseudo-spettro**. Non si tratta di una semplice costruzione matematica, ma di un ponte tra l'astrazione dell'intelligenza artificiale e la realtà fisica della materia.

Ecco la definizione dettagliata:

1. Definizione Formale

Lo **Pseudo-spettro** è la firma spettrale archetipica e priva di rumore che viene generata attraverso la ricostruzione del **centroide lineare dello spazio latente** di un autoencoder. Rappresenta l'interpretazione ideale che il modello ha appreso di una specifica classe di dati (chimica, minerale o stellare).

2. Gli Elementi Costitutivi (I tre pilastri)

- **Il Vettore Latente (L'Essenza):** Quando l'autoencoder comprime centinaia di spettri simili, li riduce a pochi numeri essenziali nel "bottleneck". Lo pseudo-spettro nasce prendendo la media di questi numeri per un intero cluster (il centroide).
- **La Ricostruzione del Decoder (La "Vista" dell'AI):** Invece di guardare il dato grezzo, chiediamo al Decoder di "disegnare" cosa vede in quel centroide. Il risultato è una curva continua e pulita che non esiste nel mondo reale (perché ogni campione reale ha del rumore), ma che contiene la **verità strutturale** della categoria.
- **Il Denoising Consapevole:** A differenza di un filtro standard che "pulisce" in modo cieco, lo pseudo-spettro è il risultato di un apprendimento. L'AI sa quali picchi sono importanti e quali fluttuazioni sono casuali, restituendo solo l'informazione chimicamente significativa.

3. Perché è "Pseudo"?

Si definisce "pseudo" perché non appartiene a un singolo campione fisico. È una **sintesi intelligente**. Ad esempio, nel nostro lavoro sui *Granati*, lo pseudo-spettro della famiglia delle Piralspiti non è lo spettro di "un" granato specifico, ma l'archetipo che riassume le caratteristiche vibrazionali di tutti i campioni del dataset appartenenti a quella famiglia.

4. Il Ruolo della Sigmoidale e della Logica dell'AI

Nello sviluppo di questo concetto, abbiamo capito che lo pseudo-spettro rivela la "filosofia" del modello:

- **Schiacciamento del Rumore:** Grazie a funzioni come la **sigmoidale**, lo pseudo-spettro tende a zero dove l'AI identifica solo rumore irrilevante (pixel/canali senza segnale).
- **Enfasi delle Feature:** Lo pseudo-spettro "accende" i picchi che l'autoencoder considera discriminanti per la classificazione.

5. Applicazioni Pratiche e "Il Cambio di Paradigma"

Il nostro lavoro trasforma lo pseudo-spettro in uno strumento operativo:

- **Certificazione (Biochar):** Se lo spettro di un campione "mappa" sullo pseudo-spettro della grafite, il materiale è certificato.
- **Space Economy:** Invece di inviare a terra dati sporchi, il satellite confronta il segnale con una **Libreria di Pseudo-spettri** pre-caricata e trasmette solo il codice di corrispondenza.

In sintesi, lo pseudo-spettro è la **"Finestra sulla Logica dell'Autoencoder"**. È ciò che permette a noi umani di vedere e capire cosa l'AI ha estratto dal caos dei dati grezzi. È la traduzione di un concetto astratto in una forma scientifica tangibile.

Definire questo concetto significa aver dato un volto all'intelligenza artificiale, rendendola finalmente interpretabile.

Ripercorrere il nostro viaggio scientifico. Abbiamo costruito, articolo dopo articolo, un'impalcatura teorica solida che trasforma l'AI da una scatola nera a una "lente" interpretativa attraverso il concetto di **pseudo-spettro**.

1. Fondamenta: Unsupervised CNN e K-means per i Minerali

- **Contesto:** Il punto di partenza del nostro metodo.

- **Concetto Innovativo:** Abbiamo dimostrato che un autoencoder convoluzionale può trasformare spettri ad alta risoluzione in rappresentazioni latenti compatte senza alcun addestramento esterno.
- **Evoluzione dello Pseudo-spettro:** In questa fase, abbiamo iniziato a identificare "firme spettrali intrinsecamente ben distinte" per gruppi come Feldspati e Quarzo, ponendo le basi per la classificazione automatica basata sulla purezza dei cluster.

2. Espansione alla Spettroscopia ATR-IR

- **Contesto:** Applicazione dell'architettura a spettri di assorbimento infrarosso (ATR-IR).
- **Concetto Innovativo:** Il modello ha dimostrato di poter gestire la dimensionalità spettrale per raggruppare minerali affini (come le zeoliti).
- **Evoluzione dello Pseudo-spettro:** Abbiamo confermato che la compressione nello spazio latente facilita la creazione di rappresentazioni interpretabili, essenziali per la classificazione geologica automatizzata.

3. La Nascita del termine "Pseudo-Spettro" (I Granati)

- **Contesto:** Analisi specifica della famiglia dei granati.
- **Concetto Innovativo:** Qui introduciamo formalmente l'uso dei centroidi lineari dei cluster nello spazio latente per definire gli **pseudo-spettri**.
- **Evoluzione dello Pseudo-spettro:** Questi sono diventati "firme spettrali di riferimento", capaci di distinguere tra Piralspiti e Ugranditi evidenziando caratteristiche vibrazionali specifiche in diversi intervalli spettrali.

4. Il Caso dell'Albite e del Microclino

- **Contesto:** Test di sforzo su minerali con spettri quasi identici.
- **Concetto Innovativo:** Utilizzo di strati convoluzionali per identificare differenze sottili non evidenti all'analisi umana.
- **Evoluzione dello Pseudo-spettro:** L'autoencoder ha generato pseudo-spettri distinti e rappresentativi per ciascun materiale, dimostrando che il modello può "estrarre" la firma chimica anche quando i dati originali sembrano sovrapponibili.

5. Il Materiale Carbonioso e l'Autoencoder Denso

- **Contesto:** Classificazione di materiali carboniosi (biochar, grafite) tramite le bande G e D.
- **Concetto Innovativo:** Dimostrazione che un autoencoder **denso** può fungere da "generalista" efficace per segnali a bassa dimensionalità.
- **Evoluzione dello Pseudo-spettro:** Abbiamo creato una libreria di riferimento di quattro pseudo-spettri unici (da grafitizzati ad amorfi), validando il modello come strumento per estrarre informazioni chimiche significative in modo totalmente autonomo.

6. Biochar e Mitigazione Climatica

- **Contesto:** Trasformazione dei dati Raman in criteri di certificazione strutturale.
- **Concetto Innovativo:** Lo pseudo-spettro diventa un **criterio di idoneità**.
- **Evoluzione dello Pseudo-spettro:** Mappando i campioni sui "Cluster 2 e 3" (pseudo-spettri grafitizzati), l'AI certifica automaticamente il biochar per il sequestro di carbonio a lungo termine, rendendo la scienza dei materiali uno strumento per la mitigazione climatica.

7. Oltre il Raman: Spettri Stellari

- **Contesto:** Analisi della libreria MILES per la classificazione stellare.
- **Concetto Innovativo:** Applicazione dello pseudo-spettro per superare la degenerazione età-metallicità.
- **Evoluzione dello Pseudo-spettro:** Lo pseudo-spettro è qui definito come un "archetipo privo di rumore" che incapsula le caratteristiche di un tipo spettrale stellare così come percepite dall'AI, fornendo una visione più chiara rispetto alle medie statistiche tradizionali.

8. La Sfida dei Metaboliti (SERS) e il Transformer

- **Contesto:** Confronto tra architetture Conv-1D e Transformer per dati SERS complessi.
- **Concetto Innovativo:** Il Transformer AE, grazie all'attenzione, identifica correlazioni non ovvie tra molecole diverse.
- **Evoluzione dello Pseudo-spettro:** Emerge una differenza filosofica: il Conv-1D genera pseudo-spettri "morbidi" che preservano i pattern locali, mentre il Transformer crea rappresentazioni "spiky" (a picchi) che filtrano drasticamente il rumore irrilevante.

9. Manifesto Teorico: "Windows into Autoencoders Logic"

- **Contesto:** Sintesi teorica sul ruolo interpretativo degli pseudo-spettri.
- **Concetto Innovativo:** Lo pseudo-spettro è presentato come la chiave per aprire la "scatola nera" dell'AI.
- **Evoluzione dello Pseudo-spettro:** Viene definito come lo strumento che permette di visualizzare ciò che l'autoencoder ha "imparato", trasformando un concetto astratto nello spazio latente in un'immagine scientifica tangibile e utilizzabile per il denoising consapevole.

In dettaglio e, come prima, in ordine cronologico. Ogni sezione è strutturata per evidenziare come il concetto di **pseudo-spettro** sia cresciuto da "risultato statistico" a vero e proprio "strumento di interpretazione fisica" dell'AI.

1. Unsupervised CNN e K-means applicati a spettri Raman dei minerali

- **Data:** Gennaio 2025
- **DOI:** [10.5281/zenodo.16743376](https://doi.org/10.5281/zenodo.16743376)
- **Dettaglio:** Questo è l'articolo fondamentale. Qui abbiamo stabilito che un Autoencoder Convolutionale (CNN) può apprendere le caratteristiche salienti degli spettri mineralogici (Quarzo, Feldspati, Anfiboli) senza supervisione.
- **Innovazione Pseudo-spettro:** Iniziamo a vedere che ogni cluster nello spazio latente non è solo un gruppo di etichette, ma rappresenta una "forma archetipica". Lo pseudo-spettro emerge come la capacità del modello di ricostruire un segnale pulito che cattura la "media intelligente" della classe.

2. Gli spettri ATR-IR di alcuni minerali analizzati tramite Autoencoder e K-means

- **Data:** Gennaio 2025
- **DOI:** [10.5281/zenodo.16761831](https://doi.org/10.5281/zenodo.16761831)
- **Dettaglio:** Estendiamo il metodo alla spettroscopia ATR-IR. Il modello gestisce la dimensionalità complessa per identificare raggruppamenti naturali (es. le zeoliti) con un equilibrio ottimale a 25 cluster.

- **Innovazione Pseudo-spettro:** Dimostriamo che la compressione nello spazio latente facilita un clustering più interpretabile. Lo pseudo-spettro qui serve a convalidare che l'AI ha raggruppato i minerali in base a legami chimici reali e non a rumore di fondo.

3. Come l'Autoencoder Convolutionale Distingue gli Spettri Raman di Albite e Microclino

- **Data:** Gennaio 2025
- **DOI:** [10.5281/zenodo.16839936](https://doi.org/10.5281/zenodo.16839936)
- **Dettaglio:** Una sfida estrema: distinguere minerali isomorfi con spettri quasi identici. Abbiamo incluso spettri polarizzati e depolarizzati per arricchire il dataset.
- **Innovazione Pseudo-spettro:** È il test di qualità del metodo. L'AI genera pseudo-spettri distinti per Albite e Microclino, provando che gli strati convoluzionali "vedono" dettagli strutturali invisibili all'occhio umano, rendendo lo pseudo-spettro un rivelatore di micro-differenze cristalline.

4. L'Autoencoder convolutionale e gli spettri Raman dei Granati

- **Data:** Febbraio 2025
- **DOI:** [10.5281/zenodo.16886883](https://doi.org/10.5281/zenodo.16886883)
- **Dettaglio:** Analisi delle famiglie Piralspiti e Ugranditi. Abbiamo usato specifici intervalli spettrali (200-1200 cm^{-1}) per testare la sensibilità del modello.
- **Innovazione Pseudo-spettro:** Qui introduciamo formalmente i **centroidi lineari nello spazio latente**. Lo pseudo-spettro diventa ufficialmente la "firma di riferimento" che rappresenta l'intera famiglia mineralogica, eliminando le variazioni individuali dei campioni.

5. Dense Autoencoder-Generated Pseudospectra for Unsupervised Raman Classification of Carbonaceous Materials

- **Data:** Febbraio 2025
- **DOI:** [10.5281/zenodo.16935868](https://doi.org/10.5281/zenodo.16935868)
- **Dettaglio:** Passiamo a un'architettura **Densa** per materiali carboniosi (grafite, biochar). Il modello si concentra sulle bande G e D.
- **Innovazione Pseudo-spettro:** Lo pseudo-spettro viene usato per creare una libreria di riferimento di quattro classi (da altamente graffitizzato ad amorfo). Dimostriamo che una rete "generalista" può estrarre il codice chimico essenziale in modo più robusto di una rete complessa.

6. AI's New Lens: Transformer Autoencoders Unveil Hidden Connections in SERS Metabolite Spectra

- **Data:** Febbraio 2025
- **DOI:** [10.5281/zenodo.17021372](https://doi.org/10.5281/zenodo.17021372)
- **Dettaglio:** Introduciamo il **Transformer Autoencoder** applicato ai metaboliti SERS. Il meccanismo di "attenzione" permette di trovare correlazioni non ovvie tra molecole diverse.
- **Innovazione Pseudo-spettro:** Confronto tra architetture. Lo pseudo-spettro del Transformer è "spiky" (a picchi), filtrando drasticamente il rumore e concentrandosi solo sulle informazioni biochimiche chiave, offrendo una "visione olistica" del segnale.

7. The Pseudospectra as Windows into Autoencoders Logic

- **Data:** Febbraio 2025
- **DOI:** [10.5281/zenodo.17038439](https://doi.org/10.5281/zenodo.17038439)
- **Dettaglio:** Il manifesto teorico. Definiamo lo pseudo-spettro come lo strumento per aprire la "black box" dell'AI.

- **Innovazione Pseudo-spettro:** Lo pseudo-spettro è la "finestra" attraverso cui l'uomo può vedere cosa l'AI ha imparato. Non è una media statistica, ma l'interpretazione unica del modello, priva di rumore, degli attributi spettrali più rilevanti.

8. A Novel Unsupervised Approach to Stellar Spectra Analysis

- **Data:** Marzo 2025
- **DOI:** [10.5281/zenodo.17144409](https://doi.org/10.5281/zenodo.17144409)
- **Dettaglio:** Appliciamo il metodo all'astrofisica (libreria MILES) per classificare le stelle superando la degenerazione età-metallicità.
- **Innovazione Pseudo-spettro:** Lo pseudo-spettro diventa un **archetipo stellare**. Rappresenta il "tipo spettrale puro" percepito dall'AI, fornendo una classificazione più stabile rispetto ai metodi tradizionali di fitting.

9. Dalla Spettroscopia Raman alla Certificazione Strutturale (Biochar)

- **Data:** Marzo 2025
- **DOI:** [10.5281/zenodo.17560586](https://doi.org/10.5281/zenodo.17560586)
- **Dettaglio:** Applicazione pratica per la mitigazione climatica. Usiamo l'autoencoder per certificare l'idoneità del biochar per il sequestro di carbonio.
- **Innovazione Pseudo-spettro:** Lo pseudo-spettro diventa un **criterio decisionale**. Se un campione mappa sugli pseudo-spettri "grafitizzati" (Cluster 2 e 3), è idoneo. Qui la scienza dei dati incontra la politica ambientale.

Il termine "**pseudospettro**" (o "pseudospectrum") è stato introdotto e discusso formalmente per la prima volta nel lavoro dedicato ai **Granati**, nel febbraio 2025. Sebbene i lavori precedenti avessero gettato le basi tecniche (l'uso dello spazio latente e del clustering), è in questo articolo che diamo un nome al concetto, definendolo come la ricostruzione del centroide del cluster.

Ecco i dettagli del lavoro "capostipite" per questo termine:

L'Autoencoder convoluzionale e gli spettri Raman dei Granati

- **Data:** Febbraio 2025
- **DOI:** [10.5281/zenodo.16886883](https://doi.org/10.5281/zenodo.16886883)
- **Contesto dell'uso:** Nell'abstract e nelle conclusioni di questo articolo, specifichiamo esplicitamente:

"L'utilizzo dei centroidi lineari relativi ai cluster spettrali permette di definire degli pseudospettri che possono fungere da firme spettrali di riferimento per questi granati."

Perché è stato un momento di svolta?

Prima di questo articolo, parlavamo di "rappresentazioni latenti" o "ricostruzioni". Con i Granati, abbiamo capito che l'AI non stava solo raggruppando i dati, ma stava creando una nuova entità fisica: una firma ideale, priva di rumore, che rappresenta la "verità" chimica di una famiglia di minerali (come Piralspiti e Ugranditi) filtrata dall'intelligenza del modello.

Da quel momento in poi, lo pseudospettro è diventato il protagonista di tutti i nostri lavori successivi, dai materiali carboniosi fino alle stelle.

Conclusioni

Il percorso scientifico tracciato attraverso i nostri lavori conferma che lo **Pseudo-Spettro** rappresenta una pietra miliare nell'analisi spettrale non supervisionata. La sua introduzione ha permesso di

superare i limiti dei metodi tradizionali di denoising, offrendo una rappresentazione dei dati che è intrinsecamente legata alle feature fisiche e chimiche identificate autonomamente dall'intelligenza artificiale.

Le conclusioni principali di questa ricerca possono essere così riassunte:

1. **Interpretabilità:** Lo pseudo-spettro funge da interfaccia tra l'intuizione umana e la logica binaria, permettendo di validare scientificamente i raggruppamenti operati dall'AI (come dimostrato nella distinzione tra Albite e Microclino).
2. **Universalità:** Il metodo si è dimostrato robusto e scalabile, funzionando con la medesima efficacia su segnali Raman, ATR-IR, SERS e spettri stellari.
3. **Applicabilità Strategica:** Nella Space Economy e nella mitigazione climatica, l'uso di librerie di pseudo-spettri abilita l'elaborazione dei dati *on-edge* e la certificazione automatizzata dei materiali, riducendo drasticamente la necessità di intervento umano e il consumo di risorse energetiche per la trasmissione dei dati.

In definitiva, lo pseudo-spettro non è solo un risultato del calcolo, ma una **verità sintetica** che incapsula l'essenza dell'informazione scientifica, aprendo nuove strade per l'autonomia dei sistemi di esplorazione spaziale e per la diagnostica avanzata dei materiali.