Università degli Studi di Padova

Dipartimento di Matematica "Tullio Levi-Civita"

Corso di Laurea in Informatica

Evoluzione Temporale della Complessità Visiva nelle Copertine Musicali: Un'Analisi Quantitativa

Tesi di Laurea Triennale

Relatore

Dott. Cecconello Stefano

Correlatore

Prof. Galeazzi Alessandro

Laureando

Nicolas Fracaro

Matricola 2076420

Anno Accademico 2024-2025

© Nicolas Fracaro, Luglio 2025. Tutti i diritti riservati. Tesi di Laurea Triennale: "*Evoluzione Temporale della Complessità Visiva nelle Copertine Musicali: Un'Analisi Quantitativa*", Università degli Studi di Padova, Dipartimento di Matematica "Tullio Levi-Civita".

Sommario

Il ruolo della copertina di un album musicale, da semplice involucro protettivo a fondamentale strumento artistico e di marketing, si è profondamente evoluto con il mutare delle tecnologie di fruizione. Partendo da questo presupposto, la presente tesi si propone di indagare su larga scala l'evoluzione temporale della complessità visiva delle copertine attraverso un'analisi quantitativa.

Per raggiungere tale obiettivo, è stato costruito e standardizzato un dataset di oltre 46.000 album. L'analisi è stata condotta tramite un approccio multi-metrico al fine di ottenere una misurazione robusta e sfaccettata della complessità.

I risultati evidenziano che, sebbene la maggior parte dei generi musicali si conformi a una "norma" stilistica comune, generi come *Metal* e *Hip Hop* si distinguono per una complessità visiva mediamente più elevata, imputabile a un maggiore disordine. L'analisi temporale ha poi rivelato una tendenza generale condivisa, caratterizzata da un diffuso calo della complessità media a partire dal 2009. Parallelamente a questa tendenza, si è però osservato un progressivo aumento della variabilità stilistica nel tempo.

Nel loro insieme, questi risultati suggeriscono che l'evoluzione del linguaggio visivo delle copertine musicali non sia casuale, ma rifletta un adattamento alle trasformazioni dell'ecosistema che le circonda.

Ringraziamenti

Desidero rivolgere un sentito ringraziamento ai miei relatori, il Dott. Cecconello Stefano e il Prof. Galeazzi Alessandro, per i loro preziosi consigli e il costante supporto durante l'esperienza di stage e la stesura di questa tesi.

Infine, un ringraziamento speciale va a mia madre, che con i suoi sacrifici mi ha dato la possibilità di intraprendere questo percorso universitario. In questo titolo di studio c'è il mio nome, ma è un traguardo che in realtà appartiene a entrambi.

Padova, Luglio 2025 *Nicolas Fracaro*

Indice

1 Introduzione **1**

1.1 Struttura della tesi 2

2 Background **3**

2.1 La Complessità Visiva di un'Immagine 3

2.2 Metriche di Complessità per Immagini 4

2.3 L'Analisi della Complessità: un Approccio Interdisciplinare . . . 5

2.4 L'Analisi delle Copertine degli Album 6

2.5 Obiettivi e Contributo della Tesi 8

3 Dataset **9**

3.1 MuMu Dataset 9

3.2 MSD-I Dataset 10

3.3 Billboard Dataset 11

3.4 Standardizzazione e Labelling del Dataset tramite LLM 12

3.4.1 Mapping dei Generi a un Set Standardizzato 13

3.4.2 Predizione dei Generi Mancanti 13

3.4.3 Predizione della Data di Pubblicazione 14

3.5 Dataset Finale 15

4 Metriche di complessità **18**

4.1 Complessità Significativa tramite MDL Clustering (MDLc) . . . 18

4.1.1 Utilizzo del Principio MDL 19

4.1.2 Processo Algoritmico Gerarchico 20

4.1.3 Implementazione Pratica 21

4.2 Piano Entropia-Complessità 22
4.2.1 Le due metriche: Entropia e Complessità 22
4.2.2 Processo di Calcolo 23
4.3 Complessità basata sulla Compressione (ZIPc) 24
4.3.1 Processo di Calcolo 25
4.4 Validazione Qualitativa delle Metriche 26
4.4.1 Analisi della Metrica MDLc 26
4.4.2 Analisi della Metrica ZIPc 28
4.4.3 Analisi del Piano Entropia-Complessità 29
4.5 Object Detection con YOLO 30
4.5.1 Implementazione Pratica 31

5 Strumenti Utilizzati **33**
5.1 Python e librerie 33
5.2 Esecuzione in Batch e Checkpointing 34
5.3 Ottimizzazione tramite Parallelismo 34

6 Risultati **35**
6.1 Analisi della Complessità per Genere Musicale 35
6.2 Evoluzione della Complessità nel Tempo 40

7 Conclusioni **47**

Bibliografia **i**

Elenco delle figure

3.1 Distribuzione temporale del numero di album provenienti dai dataset MuMu e MSD-I. 10
3.2 Distribuzione temporale del numero di album provenienti dal dataset Billboard. 12
3.3 Conteggio totale degli album per genere all'interno del dataset finale. 16
3.4 Grafico a barre impilate che illustra la composizione del dataset per anno e genere. Ogni barra verticale rappresenta il numero totale di album pubblicati in un anno, mentre i segmenti colorati al suo interno ne mostrano la suddivisione per genere, come indicato nella legenda. 17

4.1 Griglia di 9 immagini campione per l'analisi qualitativa delle metriche di complessità. 27
4.2 Complessità delle 9 immagini campione calcolata con la metrica MDLc. Valori più alti indicano una maggiore complessità significativa. 28
4.3 Complessità delle 9 immagini campione calcolata tramite il rapporto di compressione ZIP (ZIPc). Un valore più alto indica una minore comprimibilità e, quindi, una maggiore complessità secondo questa metrica. 29

4.4 Posizionamento delle 9 immagini campione nel Piano Entropia-Complessità. Il piano distingue tra disordine casuale (alta entropia, bassa complessità) e complessità strutturata (alta complessità statistica). 30

6.1 Box plot della distribuzione del punteggio di complessità MDLc per ciascuno degli 11 generi. La linea nera indica la mediana, la linea rossa tratteggiata la media. La colorazione è proporzionale al numero di album per genere. 36

6.2 Box plot della distribuzione del rapporto di compressione (ZIPc) per genere.La linea nera indica la mediana, la linea rossa tratteggiata la media. La colorazione è proporzionale al numero di album per genere. 36

6.3 Posizionamento dei valori medi di Entropia di Permutazione (asse x) e Complessità Statistica (asse y) per ciascun genere. Le barre di errore indicano l'errore standard della media (SEM). 38

6.4 Percentuale di album con oggetti (in arancione, parte inferiore) e senza oggetti (in viola, parte superiore) rilevabili da YOLOv8, suddivisi per genere musicale. 40

6.5 Andamento della complessità media (MDLc) nel tempo, suddivisa per genere. L'asse temporale è partizionato in periodi dinamici, ciascuno contenente almeno 3000 album totali (ad eccezione dell'ultimo, con 2269). Un genere viene mostrato in un dato periodo solo se sono presenti almeno 50 album. 41

6.6 Andamento della complessità media (ZIPc) nel tempo, suddivisa per genere. L'asse temporale è partizionato in periodi dinamici, ciascuno contenente almeno 3000 album totali (ad eccezione dell'ultimo, con 2269). Un genere viene mostrato in un dato periodo solo se sono presenti almeno 50 album. 42

6.7 Evoluzione dei valori medi di Entropia di Permutazione e Complessità Statistica nel tempo. I punti rappresentano periodi temporali dinamici, ciascuno contenente un minimo di 3000 album ad eccezione dell'ultimo periodo che ne comprende 2269. La linea tratteggiata collega i punti in ordine cronologico, mostrando la traiettoria evolutiva aggregata delle copertine. 43
6.8 Evoluzione nel tempo del numero medio di oggetti per album riconosciuti da YOLOv8, suddivisi per genere. L'asse temporale è partizionato in periodi dinamici, ciascuno contenente almeno 3000 album totali (ad eccezione dell'ultimo, con 2269). Un genere viene mostrato in un dato periodo solo se sono presenti almeno 50 album. 45
6.9 Box plot della distribuzione della complessità MDLc per periodo temporale. I dati sono raggruppati in periodi dinamici contenenti almeno 3000 album (eccetto l'ultimo che ne contiene 2269). . . . 46
6.10 Box plot della distribuzione del rapporto di compressione (ZIPc) per periodo temporale, con i dati raggruppati in periodi dinamici contenenti almeno 3000 album (eccetto l'ultimo che ne contiene 2269). 46

Elenco delle tabelle

3.1 Distribuzione dei generi predetti dal modello LLM con alta confidenza sugli album senza etichetta. 14
3.2 Accuratezza della predizione dell'anno di pubblicazione per le 289 risposte ad alta confidenza. 15

4.1 Elenco delle 80 classi di oggetti riconosciute da YOLOv8. 31

6.1 Occorrenze per le 79 classi di oggetti rilevate nel dataset, ordinate in modo decrescente. In totale, sono stati identificati 77.422 oggetti. 39

Capitolo 1

Introduzione

In quanto componente fondamentale di ogni società, la musica agisce come specchio e motore dei cambiamenti culturali e tradizionali [1, 2, 3]. Questo suo ruolo dinamico si è evoluto di pari passo con la tecnologia, che ha ripetutamente trasformato le pratiche di composizione e fruizione nel corso della storia [4]. In questo scenario, l'avvento della musica riproducibile ha introdotto una componente nuova e fondamentale: la sua dimensione visiva.

Dal momento in cui la musica è stata commercializzata con la diffusione dei primi vinili, è emersa la necessità di darle un volto, un'identità visiva che potesse rappresentarla prima ancora dell'ascolto. La copertina di un album è quindi diventata il primo, cruciale punto di contatto tra l'opera musicale e il pubblico. Inizialmente concepita come semplice confezione protettiva, si è rapidamente trasformata in una vera e propria tela, un'estensione dell'espressione artistica del musicista. Figure pionieristiche come Alex Steinweiss, considerato l'inventore della copertina d'album moderna, hanno compreso che un'immagine evocativa poteva non solo catturare l'essenza della musica contenuta, ma anche arricchirla, diventando parte integrante dell'atmosfera e del messaggio che l'artista intendeva trasmettere [5].

Con il passare del tempo, tuttavia, il ruolo della copertina si è evoluto, assumendo una doppia valenza. Parallelamente alla sua funzione artistica, con l'esplosione del mercato discografico e il consumo di massa, la copertina è diventata un potente strumento di marketing e branding [6]. Le opere visive hanno

iniziato una competizione per catturare l'attenzione del potenziale acquirente, prima sugli scaffali affollati dei negozi di dischi e poi, in modo ancora più intenso, negli "scaffali virtuali" delle piattaforme digitali. In questo contesto digitale, dove l'ascoltatore scorre centinaia di miniature, la capacità di una copertina di emergere in una frazione di secondo è diventata essenziale[7], influenzandone potenzialmente la progettazione.

Il cambiamento nei **paradigmi di fruizione** potrebbe quindi non influenzare solo le strategie di marketing, ma anche esercitare una potenziale pressione selettiva sulla natura stessa dei contenuti. Questa ipotesi sembra trovare un parallelo in recenti studi che hanno già osservato una tendenza misurabile verso la semplificazione in altri domini creativi, come nei testi delle canzoni [8] e nella struttura della musica stessa [9]. Diventa perciò legittimo chiedersi: se la musica e le sue parole si stanno adattando a questo ecosistema, anche la loro controparte visiva, la copertina, sta subendo un'evoluzione simile?

Partendo da questo contesto, la presente tesi si propone di indagare, da un punto di vista oggettivo e quantificabile, come la **complessità visiva** delle copertine degli album musicali si sia evoluta nel tempo e attraverso i diversi generi.

1.1 Struttura della tesi

Per rispondere a queste domande, il presente lavoro di tesi è strutturato come segue. Inizialmente, sarà fornito un inquadramento teorico, analizzando la letteratura relativa alle metriche di complessità e agli studi sull'evoluzione estetica delle copertine musicali. Successivamente, verrà descritta la metodologia adottata, illustrando nel dettaglio prima il processo di costruzione del dataset e poi le metriche di complessità selezionate dal punto di vista teorico e pratico. Si procederà quindi con l'applicazione di tali metriche, presentando e discutendo i risultati emersi dall'analisi. Il capitolo conclusivo, infine, riassumerà i principali contributi della tesi, riflettendo su possibili sviluppi futuri interessanti.

Capitolo 2

Background

Il presente capitolo fornisce il quadro concettuale necessario a comprendere l'analisi di questa tesi. Inizialmente, si definisce il concetto di **complessità visiva** e si presentano le metriche utilizzate per la sua misurazione. Successivamente, si esplora lo stato della ricerca attuale, analizzando sia l'uso della complessità come metrica interdisciplinare, sia gli studi specifici dedicati alle copertine degli album.

2.1 La Complessità Visiva di un'Immagine

La definizione di "complessità" di un'immagine rappresenta una sfida concettuale ancora aperta nella ricerca scientifica. Non esiste, infatti, una definizione universalmente accettata di complessità visiva, poiché questo concetto può essere analizzato secondo diverse prospettive teoriche e metodologiche, come evidenziato da Donderi nella sua rassegna [10].

Da un punto di vista puramente computazionale, il concetto di complessità può essere formalizzato attraverso la **complessità di Kolmogorov**. Questa definisce la complessità di un oggetto come la lunghezza del più breve programma per computer in grado di produrlo come output [11]. Nel contesto delle immagini, questo si traduce nella quantità minima di informazione necessaria per ricostruire l'immagine originale. Sebbene questa definizione sia teoricamente elegante, la sua applicazione pratica presenta un limite importante: è una funzione **non calcolabile** ma solo approssimabile. Per questo motivo si sono

sviluppate tecniche alternative che considerano fattori più direttamente legati alla percezione visiva, nel tentativo di avvicinarsi alla complessità effettivamente percepita dall'occhio umano. È importante notare che queste tecniche di analisi quantitativa non sempre riflettono fedelmente la complessità percepita dall'osservatore umano [12].

Questi approcci alternativi, orientati verso la complessità percepita, si basano su un insieme di fattori tra cui: la **variabilità a basso livello** (proprietà statistiche dei pixel come colore, intensità e texture), **l'organizzazione spaziale** (presenza di pattern, simmetrie e dettagli strutturali) e il **contenuto semantico** (numero e varietà di oggetti riconoscibili) [10].

La natura intrinsecamente multidimensionale della complessità visiva rende quindi la sua misurazione una sfida metodologica significativa, che ha portato allo sviluppo di diverse metriche computazionali, ognuna mirata a catturarne un aspetto differente.

2.2 Metriche di Complessità per Immagini

Per quantificare la complessità visiva, la letteratura scientifica ha esplorato una vasta gamma di approcci metodologici. Un primo approccio, ispirato alla teoria della complessità di Kolmogorov, utilizza il **rapporto di compressione** di un'immagine come stima pratica della sua complessità. L'idea di fondo è che un'immagine semplice, con pattern ripetitivi, sia altamente comprimibile, mentre un'immagine ricca di dettagli lo sia molto meno. Di conseguenza, una dimensione ridotta del file compresso è indicativa di bassa complessità, mentre una dimensione maggiore suggerisce una complessità visiva più elevata [13].

Un'altra famiglia di tecniche analizza l'immagine nel **dominio della frequenza**. Metodi come la **Trasformata di Fourier** e la **Trasformata Wavelet** scompongono l'immagine nelle sue frequenze costitutive. La complessità viene quindi stimata in base alla distribuzione dell'energia nello spettro: una maggiore presenza di componenti ad alta frequenza, che corrispondono a dettagli fini e variazioni rapide, è indice di un'immagine più complessa. Questi metodi sono

stati applicati con successo sia in contesti di computer vision per la medicina [14] sia nell'analisi di opere d'arte [15].

Un approccio correlato, ma più diretto all'analisi dei dettagli, si basa sul calcolo del **gradiente dell'intensità dei pixel**. Questa metrica quantifica la presenza di bordi e texture misurando la rapidità con cui cambiano i valori di colore o luminanza; Redies et al.[16] hanno impiegato questa misura per studiare le proprietà estetiche di fotografie di opere d'arte e scene naturali .

La **dimensione frattale** offre invece una prospettiva geometrica sulla complessità, misurando l'irregolarità e il livello di dettaglio di un'immagine a diverse scale. Una dimensione frattale più alta indica una struttura più intricata e "rugosa". Lam et al.[17] hanno esplorato l'efficacia di questo metodo per caratterizzare la complessità di immagini in ambito geografico e scientifico.

Infine, approcci più recenti e sofisticati si basano su principi avanzati della teoria dell'informazione. La **Minima Lunghezza di Descrizione (MDL)** permette di misurare una "complessità significativa", definita come la quantità di informazione necessaria a descrivere l'immagine una volta segmentata in regioni coerenti [18]. Su una linea simile, il **piano Entropia-Complessità** offre una visione più ricca, posizionando l'immagine in uno spazio bidimensionale che distingue tra disordine (Entropia) e struttura (Complessità Statistica). Questo potente strumento è stato utilizzato per analizzare l'evoluzione stilistica nella storia dell'arte, tracciando le traiettorie delle opere nel tempo [19].

2.3 L'Analisi della Complessità: un Approccio Interdisciplinare

L'analisi della complessità come metrica per comprendere l'evoluzione dei sistemi costituisce un paradigma di ricerca trasversale, applicato con successo in domini che spaziano dalle scienze naturali a quelle culturali.

In biologia, ad esempio, la ricerca ha indagato come il grado di sofisticazione degli organismi possa aumentare attraverso processi evolutivi non adattivi, come mutazioni e deriva genetica [20]. Nel campo degli studi urbani, invece, si è

osservato come le città evolvano quali sistemi adattivi, dove da regole locali e interazioni semplici emergono assetti globali ordinati e strutturati, spesso di natura frattale [21].

Applicando questa stessa lente di analisi ai sistemi di comunicazione umana, la ricerca ha iniziato a svelare pattern evolutivi di interesse. Uno studio sul linguaggio utilizzato sulle piattaforme di social media, per esempio, ha evidenziato un progressivo processo di semplificazione nel tempo, suggerendo che le logiche della comunicazione digitale, veloci e basate sull'economia dell'attenzione, possano favorire forme espressive più immediate [22].

Anche nel dominio dei prodotti culturali, questo metodo di indagine si è dimostrato efficace. Studi come quello di Sigaki et al. hanno misurato la complessità di migliaia di dipinti, dimostrando come sia possibile tracciare l'evoluzione degli stili artistici nel corso della storia [19]. Per quanto riguarda la musica, Di Marco et al. (2025) hanno applicato strumenti della Network Science per rappresentare le composizioni musicali come reti, riscontrando un trend generale verso la semplificazione armonica e melodica [9]. Un calo della ricchezza lessicale e un aumento della ripetitività sono stati documentati anche per i testi delle canzoni da diversi studi indipendenti [8, 23].

Questo approccio interdisciplinare si dimostra quindi una lente d'ingrandimento versatile, capace di investigare l'evoluzione di sistemi molto diversi.

2.4 L'Analisi delle Copertine degli Album

L'analisi delle copertine degli album musicali è un campo di studio che è stato affrontato con diversi approcci metodologici. La ricerca esistente ha esplorato questi artefatti visivi sia attraverso l'applicazione di tecniche computazionali per analizzare le loro caratteristiche oggettive, sia investigando il loro impatto sulla percezione umana.

Una parte significativa della ricerca si è concentrata sull'utilizzo delle caratteristiche visive per compiti di **classificazione e predizione del genere musicale**. I primi lavori, come quello di Lıbeks e Turnbull, hanno utilizzato feature a basso livello come il colore, la texture e la presenza di volti per ad-

destrare classificatori [24]. Approcci successivi hanno sfruttato tecniche di deep learning: Lee e Baraldi, ad esempio, hanno impiegato reti neurali convoluzionali (CNN) pre-addestrate per estrarre feature di alto livello, ottenendo un miglioramento significativo nell'accuratezza della predizione del genere [25]. Altri studi hanno confermato l'esistenza di una forte correlazione tra specifiche proprietà visive (come luminosità e tonalità media) e il genere musicale [26]. Più di recente, Greenfield e Paintsil (2024) hanno affrontato il problema dei dataset sbilanciati, che aveva limitato il successo degli approcci precedenti. Utilizzando un dataset bilanciato con 18 generi, hanno confrontato le performance di due architetture moderne: DenseNet-201 e Vision Transformer (ViT16/B), entrambe pre-addestrate su ImageNet, ottenendo un miglioramento significativo nell'accuratezza rispetto ai metodi precedenti [27].

Andando oltre la classificazione, un altro filone di studi, di natura più sperimentale, ha investigato come **specifici elementi di design** influenzino la percezione e le aspettative dell'ascoltatore. Un esempio notevole è lo studio di Venkatesan et al., che ha dimostrato sperimentalmente l'esistenza di un legame tra la tipografia utilizzata e la musica attesa: caratteri più spigolosi e netti vengono associati a musica più energica e "dura" (es. heavy metal), mentre caratteri più arrotondati e morbidi generano l'aspettativa di una musica più calma e dolce (es. musica classica o pop) [28].

Infine, un esempio significativo di applicazione di metriche di complessità alle copertine degli album è lo studio di Grebenkina et al., che hanno investigato la relazione tra la complessità strutturale di un'immagine e la preferenza estetica [29]. Nello specifico, hanno utilizzato come metrica l'**entropia dell'orientamento dei bordi** (*edge-orientation entropy*), la quale misura la diversità degli angoli dei contorni presenti in un'immagine. Confrontando questa metrica con i giudizi di preferenza espressi da soggetti umani su diverse categorie di immagini (incluse le copertine degli album), gli autori hanno scoperto una correlazione positiva. È emerso, infatti, che le immagini con una maggiore varietà di orientamenti, e di conseguenza un'entropia più alta, tendono a essere giudicate come esteticamente più piacevoli. Questo studio dimostra che una misura di complessità strutturale, calcolabile automaticamente, può catturare un aspetto

importante della percezione umana.

2.5 Obiettivi e Contributo della Tesi

Come evidenziato dalla rassegna della letteratura, l'analisi quantitativa delle copertine degli album si è finora concentrata sull'uso di feature visive specifiche, come il colore o la tipografia, per compiti di classificazione del genere o per studiare limitati effetti percettivi. Tuttavia, una dimensione fondamentale rimane quantitativamente inesplorata: l'**evoluzione temporale** delle copertine. Queste, in quanto artefatti culturali, sono intrinsecamente legate ai contesti storici e alle evoluzioni tecnologiche, sociali e artistiche della loro epoca. Inoltre, in questo specifico dominio non è mai stata approfondita la nozione di **complessità visiva** come metrica di analisi.

Il contributo di questa tesi è proprio quello di tentare di colmare tale lacuna, conducendo un'analisi quantitativa su larga scala dell'evoluzione della complessità visiva nelle copertine degli album. L'obiettivo è individuare possibili tendenze nel tempo e differenze significative tra i vari generi musicali.

Capitolo 3

Dataset

In questa sezione viene descritto il processo di costruzione del dataset utilizzato per questo lavoro di tesi. Si analizzano le sue principali statistiche, l'origine dei dati e gli strumenti impiegati per la sua aggregazione e arricchimento. Il dataset finale è il risultato della fusione e dell'elaborazione di tre fonti di dati eterogenee: il MuMu dataset, l'MSD-I dataset e un corpus di dati estratto dalle classifiche di Billboard.

3.1 MuMu Dataset

Il MuMu dataset [30] è un corpus multimodale e multi-genere che integra diverse dimensioni informative per ogni brano musicale: audio, immagini di copertina, testi e metadati. Questo dataset nasce dalla combinazione di informazioni provenienti da *Amazon Reviews* e dal *Million Song Dataset* (MSD) [31].

A ogni brano è associato l'identificativo MusicBrainz [32] dell'album di appartenenza. Partendo da questa informazione, sono stati identificati gli album unici e, successivamente, sono state utilizzate le API di MusicBrainz per arricchire i metadati esistenti. In particolare, sono state recuperate due informazioni chiave: il nome dell'etichetta discografica (*record label*) e l'anno della prima pubblicazione ufficiale del disco, che in alcuni casi differiva da quello presente nel dataset originale. Dopo la rimozione degli album in cui non era presente

la data di pubblicazione il processo ha prodotto un insieme di **30.594 album unici**.

3.2 MSD-I Dataset

L'MSD-I (*Million Song Dataset - Images*) [33] nasce come un'estensione del *Million Song Dataset* [31], arricchendolo con il genere musicale e l'immagine di copertina per ciascun brano. Dal corpus di brani sono stati estratti gli album unici e sono state rimosse le occorrenze già presenti nel dataset MuMu, utilizzando l'identificativo di traccia del MSD come chiave di confronto.

Al fine di allineare i metadati con quelli del dataset MuMu, sono state nuovamente impiegate le API di MusicBrainz per integrare le informazioni mancanti, come l'anno di prima pubblicazione e il nome dell'etichetta discografica. Per associare i brani del MSD-I agli identificativi degli album su MusicBrainz, è stata utilizzata una mappa di corrispondenza specifica [34]. L'operazione ha permesso di ottenere **4.768 album unici**, arricchiti con le informazioni recuperate.

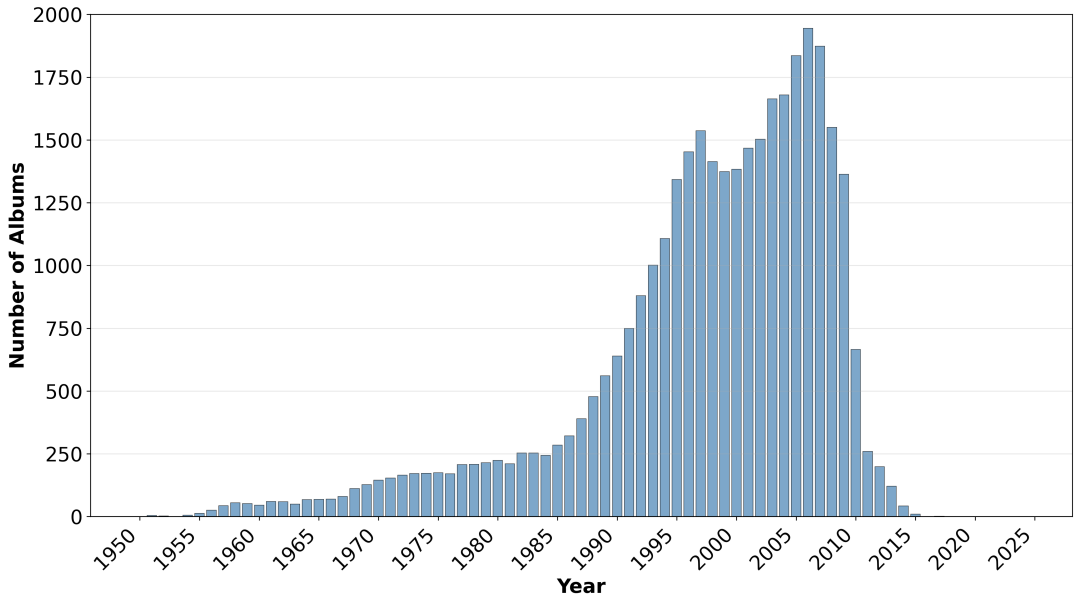


Figura 3.1: Distribuzione temporale del numero di album provenienti dai dataset MuMu e MSD-I.

3.3 Billboard Dataset

Per ovviare alla scarsa rappresentazione di album recenti nei due dataset precedenti, come si può vedere dalla figura **??**, si è deciso di integrare il corpus con dati estratti dalle classifiche di Billboard, in virtù della loro rilevanza nel mercato musicale. Sono state scaricate, con cadenza mensile da gennaio 2002 a gennaio 2025, le classifiche relative a venti diverse categorie[1]. Per l'estrazione è stata utilizzata la libreria Python `billboard.py` [35], che funge da scraper per il sito ufficiale `billboard.com`.

A seguito di una fase di pulizia, in cui sono stati rimossi i duplicati (per coppia artista-titolo) e gli album già inclusi negli altri dataset, si è ottenuto un insieme iniziale di **21.748 album unici**. I dati di partenza forniti da Billboard, limitati a titolo, artista e URL della copertina, sono stati quindi arricchiti interrogando le API di MusicBrainz. Tramite una ricerca testuale sono stati recuperati metadati essenziali come genere, tag, anno di prima pubblicazione, etichetta discografica e un URL alternativo per la copertina.

Per **16.984** di questi album è stata stabilita una corrispondenza con il database di MusicBrainz. Tuttavia, il processo di download delle relative copertine è stato ostacolato dalla non reperibilità di alcune immagini, dovuta a URL mancanti o non più validi. Eseguita quindi una rimozione di tali voci e un'ulteriore procedura di deduplicazione con l'altra parte del dataset basata questa volta sull'identificativo di MusicBrainz, il totale degli album finale si è consolidato in **11.305** album con immagine associata.

[1] Le categorie includono: `comedy`, `world`, `new-age`, `reggae`, `christian`, `tropical`, `traditional-jazz`, `rock`, `regional-mexican`, `latin-rhythm`, `latin-pop`, `independent`, `compilation`, `classical-crossover`, `classical`, `cast`, `americana-folk`, `billboard-200`, `vinyl`, `tastemaker`.

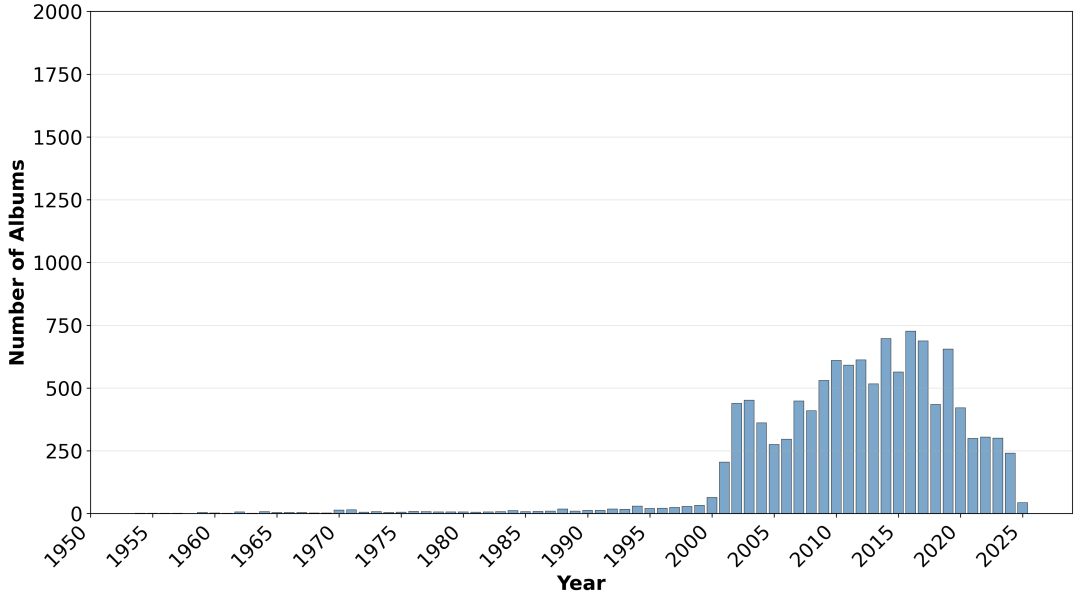


Figura 3.2: Distribuzione temporale del numero di album provenienti dal dataset Billboard.

3.4 Standardizzazione e Labelling del Dataset tramite LLM

Il dataset aggregato, pur essendo ricco di dati, presentava criticità significative che ne precludevano l'impiego diretto. Il primo problema risiedeva nell'annotazione dei generi musicali, che soffriva di un'elevata eterogeneità: erano presenti oltre 1000 etichette distinte, tag non pertinenti (es. "live", "modern creative") e una generale mancanza di standardizzazione. Il secondo problema era l'incompletezza del corpus, poiché 911 non riportavano la data di pubblicazione e 4.403 album del sotto-dataset Billboard erano privi di genere.

Per risolvere queste problematiche, è stato adottato un approccio euristico in due fasi, basato sull'impiego di Modelli Linguistici di Grandi Dimensioni (LLM), volto a standardizzare le etichette esistenti e a predire quelle mancanti.

3.4.1 Mapping dei Generi a un Set Standardizzato

La prima sfida è stata la gestione dell'elevato numero di etichette di genere. Per ricondurre questa varietà a un insieme gestibile e coerente, sono stati definiti **11 "super-generi"** principali[2]. La definizione di queste macro-categorie è stata supportata dalla consultazione della tassonomia musicale offerta da *Musicmap* [36].

Successivamente, è stato impiegato un LLM (Google Gemini 2.5 Pro, versione di Maggio 2025) [37] per eseguire due compiti:

1. Mappare ciascuna delle oltre 1000 etichette di genere esistenti a uno degli 11 super-generi.

2. Identificare e rimuovere i tag non direttamente correlati al genere musicale.

Il risultato di questo processo automatico è stato sottoposto a una **verifica manuale** per correggere eventuali errori e garantire l'accuratezza della mappa finale.

3.4.2 Predizione dei Generi Mancanti

Una parte considerevole del dataset Billboard (**4.403 album** su 16.984) era priva di etichette di genere a causa di informazioni assenti nel database di MusicBrainz. Per completare questi dati, si è ricorso a un LLM (`gemini-2.0-flash`) [38], interrogato tramite le API di Google AI.

Al modello è stato fornito in input il nome dell'artista, il titolo dell'album, la data di pubblicazione e l'ID MusicBrainz, con il compito di predire un singolo genere tra gli 11 super-generi e di associare un **livello di confidenza booleano** alla sua predizione.

Per validare l'efficacia di questa tecnica, è stato condotto un test su un campione di 500 album estratti casualmente dal dataset Billboard di cui si conosceva già il genere. I risultati sono stati i seguenti:

[2] I generi includono: Pop, Rock, Speciality, Jazz & Blues, Electronic, World Music, Country & Folk, Metal, R&B, Hip Hop, Classical.

- Su 450 album predetti con confidenza **alta ("true")**, l'accuratezza nel predire almeno uno dei generi corretti dell'album ("contained accuracy") è stata dell'**87%**.
- Sui restanti 50 album con confidenza **bassa ("false")**, la stessa metrica di accuratezza è scesa al **62%**.

Dato il buon risultato per le predizioni ad alta confidenza, si è deciso di applicare la tecnica ai dati mancanti, accettando solo le predizioni marcate come "true". Questo ha permesso di aggiungere 4.008 nuove etichette di genere valide che possiamo vedere nella tabella 3.1. Di queste, **2.510** sono state utilizzate nel dataset finale di Billboard, composto da 11.305 album.

Genere	**Numero di Album**
World Music	1071
Pop	460
Country & Folk	427
Hip Hop	420
Speciality	420
Classical	321
Rock	302
Jazz & Blues	250
R&B	227
Electronic	60
Metal	50

Tabella 3.1: Distribuzione dei generi predetti dal modello LLM con alta confidenza sugli album senza etichetta.

3.4.3 Predizione della Data di Pubblicazione

Un'ultima criticità riscontrata nel dataset era l'assenza della data di pubblicazione per un totale di **911 album**. Per tentare di recuperare questa informazione essenziale, è stato esplorato un approccio basato sull'interrogazione di un LLM (`gemini-2.0-flash`), sfruttando le API di Google AI. In particolare, è stata utilizzata una funzionalità avanzata che permette al modello di integrare i risultati di una ricerca Google per formulare la risposta.

Al modello è stato chiesto di predire l'anno della prima pubblicazione fornendo come input il nome dell'artista, il titolo dell'album e l'ID MusicBrainz, e di restituire anche un livello di confidenza booleano sulla predizione.

Per valutare l'efficacia del metodo, è stato eseguito un test su un campione di 450 album con data nota. I risultati hanno mostrato una bassa propensione del modello a restituire risposte sicure: solo **289 predizioni (64%)** sono state marcate con un livello di confidenza alto. L'accuratezza, calcolata su questo sottoinsieme di 289 risposte, è riportata nella tabella 3.2.

Metrica di Accuratezza	**Punteggio**
Corrispondenza esatta (stesso anno)	0.644
Entro 1 anno di differenza	0.716
Entro 5 anni di differenza	0.803
Stessa decade	0.782

Tabella 3.2: Accuratezza della predizione dell'anno di pubblicazione per le 289 risposte ad alta confidenza.

Considerando sia il basso tasso di risposte ad alta confidenza, sia un'accuratezza non sufficientemente elevata, **si è deciso di non adottare questa tecnica** di labelling. L'approccio finale è stato quindi quello di scartare dal dataset gli album privi della data di pubblicazione, per garantire la massima qualità e affidabilità dei dati utilizzati nell'analisi.

3.5 Dataset Finale

Al termine dei processi di pulizia, etichettatura e fusione, il dataset finale è composto da **46.399 album unici**, con data di pubblicazione compresa tra il 1950 e il 2025. È importante notare che un piccolo numero di immagini è stato escluso, in quanto ha generato un errore durante l'estrazione della metrica di complessità, impedendone il calcolo. Ogni album è etichettato con uno o più degli 11 super-generi definiti. La natura multi-label del corpus implica che un singolo album possa contribuire a più categorie di genere contemporaneamente.

La distribuzione dei generi all'interno del dataset, illustrata in Figura 3.3, risulta eterogenea. Si osserva una netta predominanza dei generi **Pop** e **Rock**, un dato che riflette la loro popolarità e vasta produzione nel mercato musicale

degli ultimi decenni. All'estremo opposto, generi come la musica **Classica** costituiscono una porzione minoritaria (1.4%), evidenziando un orientamento del dataset verso la musica popolare contemporanea.

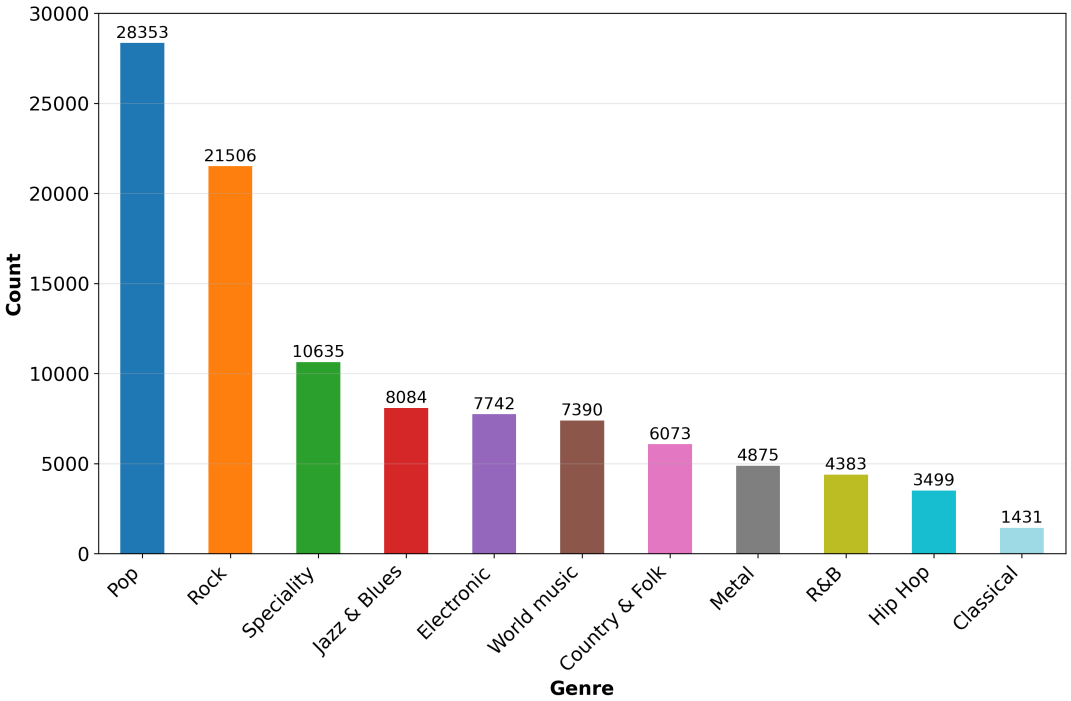


Figura 3.3: Conteggio totale degli album per genere all'interno del dataset finale.

L'analisi della distribuzione temporale, visualizzata nell'istogramma di Figura 3.4, rivela ulteriori caratteristiche chiave. Il corpus di dati è fortemente concentrato nel periodo compreso **tra il 1990 e il 2010**, che rappresenta l'intervallo di massima densità e affidabilità. La copertura per i periodi antecedenti agli anni '80 e successivi al 2015 è notevolmente inferiore, un fattore attribuibile a un **bias di campionamento** delle fonti dati utilizzate. Queste caratteristiche di non uniformità, sia a livello di genere sia temporale, dovranno essere tenute in considerazione in tutte le fasi successive di analisi.

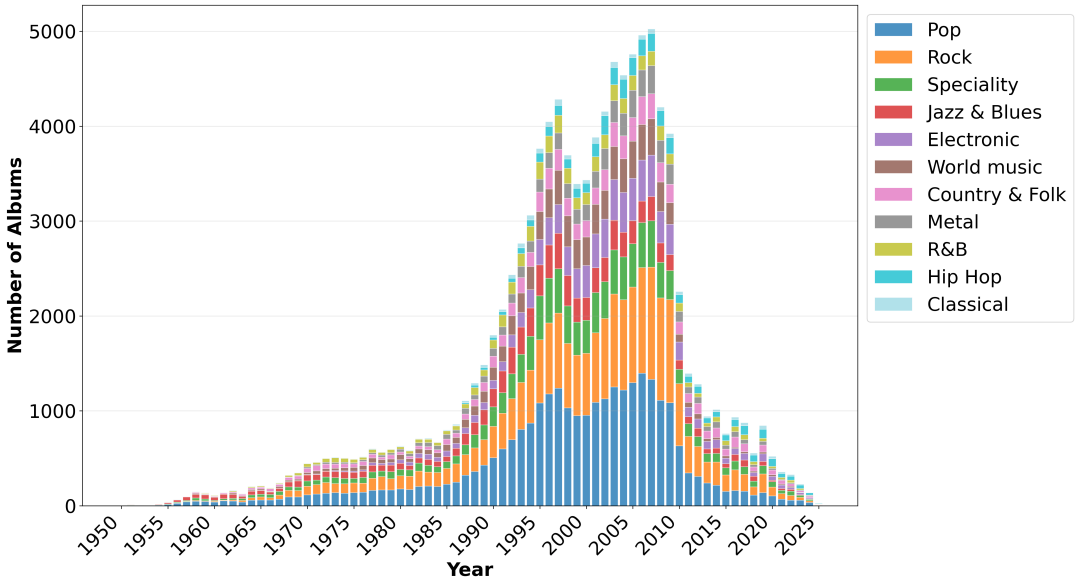


Figura 3.4: Grafico a barre impilate che illustra la composizione del dataset per anno e genere. Ogni barra verticale rappresenta il numero totale di album pubblicati in un anno, mentre i segmenti colorati al suo interno ne mostrano la suddivisione per genere, come indicato nella legenda.

Capitolo 4

Metriche di complessità

Questa sezione approfondirà nel dettaglio le metriche di complessità scelte e la loro implementazione pratica.

Non esistendo una singola definizione onnicomprensiva di "complessità", si è scelto di adottare tre approcci distinti per ottenere un'analisi sfaccettata. La metrica **MDLc** è stata selezionata per la sua capacità di misurare la complessità "significativa" utilizzando delle tecniche avanzate di clustering. Il **Piano Entropia-Complessità** offre una visione bidimensionale che separa il disordine (entropia) dalla ricchezza strutturale (complessità statistica). Infine, la metrica **ZIPc**, basata sulla teoria dell'informazione, anche se semplice, fornisce un solido punto di riferimento per il confronto.

4.1 Complessità Significativa tramite MDL Clustering (MDLc)

La metodologia per la misurazione della complessità delle immagini, descritta in questa sezione, si basa interamente sul lavoro proposto da Mahon e Lukasiewicz [18]. L'obiettivo di questo approccio è superare un limite fondamentale delle metodologie tradizionali (come il rapporto di compressione o l'entropia dei pixel), le quali tendono ad assegnare un punteggio di complessità elevato a immagini di puro rumore, che non contengono alcuna informazione significativa.

La metrica MDLc, al contrario, è progettata per quantificare la **complessità significativa**, distinguendo il segnale strutturato dal rumore casuale.

Il metodo si fonda su due principi chiave: un'applicazione pratica del principio della **Minima Lunghezza di Descrizione (MDL)** e un'**analisi gerarchica** dell'immagine per catturare la complessità a scale multiple.

4.1.1 Utilizzo del Principio MDL

Il principio MDL, derivato dalla teoria dell'informazione, afferma che il miglior modello per descrivere un insieme di dati è quello che minimizza la lunghezza totale della descrizione, ovvero quello che permette la massima compressione dei dati stessi. Nel contesto di questo algoritmo, l'MDL viene applicato in due modi pratici e sequenziali per ogni livello dell'analisi gerarchica.

Per ogni punto dell'immagine (o patch signature nei livelli successivi), l'algoritmo valuta due modi alternativi per "descriverlo" in termini di bit:

- **Descrizione tramite il Modello di Clustering:** Il costo per descrivere un punto usando il modello a mistura di gaussiane (GMM) è la somma di due componenti: il costo per indicare a quale cluster appartiene (che è basso per cluster molto popolati) e il costo per specificare la sua esatta posizione rispetto al centro del cluster (il cosiddetto "errore residuo").

- **Descrizione Diretta (come Outlier):** Il costo per descrivere un punto "da zero", senza fare affidamento sul modello, viene calcolato assumendo che il punto possa trovarsi in una qualsiasi posizione all'interno del range di valori possibili dell'immagine. Questo equivale a usare una distribuzione a priori uniforme.

Il principio MDL stabilisce che per ogni singolo punto venga scelta la descrizione più corta. Un punto casuale (rumore) difficilmente si adatterà bene a un cluster strutturato; di conseguenza, il suo errore residuo sarà alto e la sua descrizione tramite il modello risulterà "costosa". Sarà quindi più efficiente descriverlo direttamente, etichettandolo come **outlier**. Questo è il meccanismo fondamentale con cui il metodo ignora il rumore.

Una volta stabilito come calcolare il costo di descrizione per ogni punto, l'algoritmo deve scegliere il numero ottimale di cluster K da utilizzare. Anche questa scelta è guidata dall'MDL, minimizzando una funzione di costo totale:

$$\text{Lunghezza Totale}(K) = \text{Costo}_{\text{Dati}}(K) + \text{Costo}_{\text{Modello}}(K)$$

dove:

- $\mathbf{Costo_{Dati}}(K)$ è la somma delle lunghezze di descrizione più brevi (calcolate al punto precedente) per tutti i punti dell'immagine, usando un modello con K cluster.
- $\mathbf{Costo_{Modello}}(K)$ è il numero di bit necessari per memorizzare i parametri del modello stesso (medie e covarianze dei K cluster).

Si manifesta qui un trade-off cruciale: aumentare K permette di adattarsi meglio ai dati, riducendo il $\text{Costo}_{\text{Dati}}$, ma al prezzo di un $\text{Costo}_{\text{Modello}}$ più elevato. L'algoritmo testa un range di valori per K (es. da 1 a 8) e seleziona il K^* che minimizza la somma dei due costi, trovando così il modello più parsimonioso ed efficiente. Per un'immagine di rumore bianco, ad esempio, l'MDL sceglierà quasi sempre $K = 1$, poiché creare cluster aggiuntivi non giustifica il costo di descrivere il nuovo modello.

4.1.2 Processo Algoritmico Gerarchico

Il meccanismo MDL sopra descritto è il motore di un processo gerarchico che analizza l'immagine a più scale per catturare sia i dettagli locali sia le strutture globali.

- **Livello 1 (Clustering dei Pixel):** I valori dei pixel dell'immagine vengono raggruppati tramite un GMM, con K^* e outlier determinati da MDL. L'immagine viene così trasformata in una mappa 2D di indici di cluster.
- **Creazione delle "Patch Signature":** La mappa di indici viene suddivisa in piccole patch (es. 4×4) sovrapposte. Per ogni patch si calcola

una "firma" (*signature*), ovvero un vettore che conta la frequenza di ogni indice di cluster al suo interno.

- **Calcolo dell'Entropia:** Viene calcolata l'entropia di Shannon della distribuzione di tutte le firme uniche S ottenute. Questo valore, misurato in bit, rappresenta la complessità a quella specifica scala. La formula utilizzata è:

  $$H(S) = -\sum_{s \in S_{\text{unici}}} p(s) \log_2 p(s)$$

  dove $p(s)$ è la probabilità empirica (frequenza relativa) di una specifica firma s all'interno dell'insieme di tutte le firme.

- **Processo Iterativo:** Il processo viene ripetuto per i livelli successivi. L'insieme delle firme del livello i diventa l'input per il clustering del livello $i + 1$. Ad ogni iterazione, la dimensione delle patch aumenta (es. $8 \times 8, 16 \times 16, \ldots$) per catturare strutture via via più ampie.

- **Punteggio Finale:** Il punteggio di complessità totale dell'immagine è la somma delle entropie calcolate a ogni livello della gerarchia.

In sintesi, l'approccio MDLc è stato scelto per la sua capacità di discernere la struttura dal caso. Analizzando l'immagine a più scale e utilizzando l'MDL per scartare il rumore e selezionare il modello più efficiente, la metrica quantifica la complessità "significativa" in modo più profondo e robusto rispetto agli approcci tradizionali.

4.1.3 Implementazione Pratica

L'implementazione pratica della metrica MDLc si basa sul codice reso pubblicamente disponibile dagli autori Louis Mahon e Thomas Lukasiewic. Come evidenziato nel loro lavoro, la metrica presenta un costo computazionale significativo, dovuto principalmente alla fase iterativa di clustering. Per rendere l'analisi fattibile sull'intero dataset, sono state quindi adottate alcune strategie di ottimizzazione.

In linea con i suggerimenti degli autori stessi, il **numero massimo di cluster ($\mathbf{K_{max}}$)** da testare durante la ricerca MDL è stato ridotto da 8 a 5. Questa modifica, come riportato nell'articolo, permette di ridurre sensibilmente i tempi di calcolo senza un impatto apprezzabile sull'accuratezza finale della metrica.

Un'ulteriore ottimizzazione ha riguardato la standardizzazione delle immagini a una risoluzione di **224x224 pixel**. Questa scelta non solo ha ridotto i tempi di elaborazione, ma ha anche garantito la coerenza con la metodologia adottata nell'articolo di riferimento. Tutti gli altri iperparametri dell'algoritmo sono stati mantenuti ai valori di default proposti dagli autori. L'intero processo è stato infine integrato in una pipeline di elaborazione parallela, i cui dettagli tecnici sono descritti nella Sezione 5.

4.2 Piano Entropia-Complessità

Un ulteriore metodo utilizzato per caratterizzare le immagini è stata la metodologia del **Piano Entropia-Complessità**, basata sul lavoro di Sigaki et al. [19]. Questo approccio non si limita a una singola misura, ma posiziona ogni immagine in uno spazio bidimensionale definito da due metriche complementari: l'Entropia di Permutazione (H) e la Complessità Statistica (C). Insieme, queste due misure quantificano il grado di ordine e la ricchezza strutturale dei pattern spaziali locali presenti in un'immagine.

4.2.1 Le due metriche: Entropia e Complessità

- **Entropia di Permutazione (H):** Questa metrica misura il grado di **disordine** o casualità nei pattern locali formati dai pixel. Un valore di **H vicino a 0** indica un'immagine molto ordinata e ripetitiva (come un'opera minimalista), dove pochi pattern di pixel si ripetono costantemente. Al contrario, un valore di **H vicino a 1** indica un'immagine molto disordinata, dove i valori dei pixel vicini appaiono in modo quasi casuale (come il "rumore bianco" di un televisore senza segnale).

- **Complessità Statistica (C):** Questa metrica quantifica la **complessità strutturale**. A differenza dell'entropia, la complessità non è massimale per il disordine. Il suo valore è **vicino a 0** per entrambi gli estremi: sia per il perfetto ordine (struttura banale e ripetitiva) sia per il disordine totale (assenza di struttura). La complessità è invece **maggiore di 0** quando un'immagine presenta pattern spaziali elaborati, che sono ricchi di informazione ma non puramente casuali.

4.2.2 Processo di Calcolo

Il calcolo di H e C si basa sull'analisi dei **pattern ordinali** a livello locale. Sebbene la derivazione formale sia complessa, il processo può essere riassunto come segue:

1. L'immagine viene prima convertita in una matrice a scala di grigi.

2. L'algoritmo analizza l'immagine attraverso una piccola "finestra" mobile (nel paper di riferimento come nell'implentazione utilizzata in questa tesi, di dimensioni 2x2 pixel).

3. Per ogni finestra, non si considerano i valori assoluti dei pixel, ma il loro **ordine relativo**. Ad esempio, si determina quale dei 4 pixel è il più scuro, quale il secondo più scuro, e così via. Questa relazione d'ordine è chiamata "pattern ordinale".

4. Si calcola la frequenza relativa di ciascuno dei possibili pattern ordinali (24 per una finestra 2x2) su tutta l'immagine, ottenendo una distribuzione di probabilità P.

5. Da questa distribuzione di probabilità $P = \{p_i\}_{i=1,\dots,n}$, dove $n = (d_x d_y)!$ è il numero totale di possibili pattern, vengono infine calcolate le due metriche. L'**Entropia di Permutazione Normalizzata (H)**, che misura il grado di disordine, è definita come l'entropia di Shannon normalizzata

dal suo valore massimo:

$$H(P) = -\frac{1}{\ln(n)} \sum_{i=1}^{n} p_i \ln(p_i) \tag{4.1}$$

La **Complessità Statistica (C)**, che misura la presenza di strutture, è calcolata come:

$$C(P) = \frac{D_{JS}(P,\mathcal{U}) \cdot H(P)}{D_{\max}} \tag{4.2}$$

dove $\mathcal{U}$ è la distribuzione uniforme, $D_{JS}(P,\mathcal{U})$ è la **Divergenza di Jensen-Shannon** (una misura di "distanza" tra la distribuzione osservata P e una distribuzione puramente casuale $\mathcal{U}$), e $D_{\max}$ è una costante di normalizzazione.

Posizionando ogni immagine come un punto su un grafico con H sull'asse delle ascisse e C sull'asse delle ordinate, si ottiene il **Piano Entropia-Complessità**. Questo piano funge da spazio caratteristico in cui è possibile confrontare, raggruppare e analizzare diversi stili visivi in base alle loro proprietà strutturali intrinseche.

4.3 Complessità basata sulla Compressione (ZIPc)

Come terza metrica di complessità, è stato implementato un approccio basato sulla comprimibilità dei dati dell'immagine, una tecnica che trova le sue radici nella **teoria dell'informazione** e nel concetto di Complessità di Kolmogorov. L'idea fondamentale è che la complessità di un oggetto (in questo caso, un'immagine) può essere approssimata dalla sua **incomprimibilità** utilizzando un algoritmo di compressione standard [13].

In questo contesto:

- Un'immagine **semplice**, con ampie aree di colore uniforme o pattern ripetitivi, è altamente comprimibile. L'algoritmo può descriverla in modo molto efficiente, risultando in un file compresso di piccole dimensioni e, di conseguenza, in un basso rapporto di compressione.

- Un'immagine **complessa**, con molti dettagli, texture intricate e poche ripetizioni (più simile a rumore), è difficilmente comprimibile. Il file compresso avrà una dimensione vicina a quella originale, risultando in un alto rapporto di compressione.

Il punteggio di complessità, che chiameremo **ZIPc**, è quindi definito direttamente come il rapporto di compressione dell'immagine.

4.3.1 Processo di Calcolo

L'implementazione pratica di questa metrica segue una pipeline precisa per garantire la coerenza e la riproducibilità dei risultati:

1. **Standardizzazione dell'Immagine:** Per eliminare l'influenza dei formati di compressione originali (come JPEG o PNG), ogni immagine viene prima aperta e convertita in un formato standard non compresso: una mappa di pixel grezza (*raw bitmap*) con canali RGB. Questo passaggio crea una base di partenza uniforme per il calcolo della dimensione originale.

2. **Compressione:** I dati grezzi della bitmap vengono compressi utilizzando l'algoritmo **DEFLATE**, lo standard del formato ZIP, impostando il livello di compressione al massimo (`compresslevel=9`). A livello concettuale, questo algoritmo opera trovando sequenze di dati ripetute e sostituendole con brevi puntatori. Successivamente, il risultato (composto da dati non ripetuti e puntatori) viene ulteriormente compresso tramite una codifica statistica (Huffman coding) che assegna codici più corti ai simboli più frequenti.

3. **Calcolo del Rapporto di Compressione:** Il rapporto viene calcolato come:

$$\text{Rapporto di Compressione} = \frac{\text{Dimensione del file compresso}}{\text{Dimensione dei dati bitmap originali}}$$

   Un valore vicino a 0 indica alta comprimibilità (bassa complessità), mentre un valore vicino a 1 indica bassa comprimibilità (alta complessità).

4. **Punteggio di Complessità:** Il punteggio finale ZIPc è definito direttamente come il rapporto di compressione calcolato.

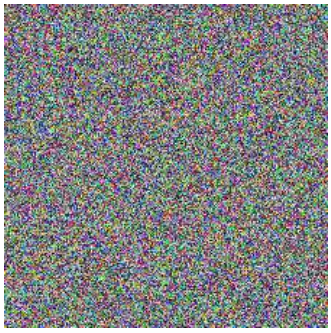
Questo metodo offre una misura di complessità oggettiva e facilmente calcolabile. Tuttavia, è importante notare che, essendo basato sulla sola intricacy statistica, tende ad assegnare un punteggio elevato anche a immagini di rumore casuale, un comportamento che si differenzia da quello della metrica MDLc.

4.4 Validazione Qualitativa delle Metriche

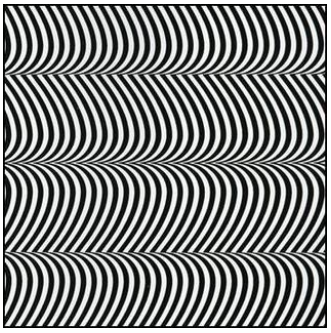
Per fornire un'interpretazione intuitiva e una validazione qualitativa delle metriche introdotte, è stata condotta un'analisi su un insieme curato di immagini campione. Questo set, visualizzato in Figura 4.1, è stato costruito per includere sia casi sintetici estremi (rumore casuale), sia una selezione di copertine di album reali con caratteristiche visive eterogenee. Lo scopo è osservare il comportamento di ogni metrica e confrontarne i punti di forza e le diverse "interpretazioni" del concetto di complessità rimanendo nel dominio dei dati reali del nostro studio.

4.4.1 Analisi della Metrica MDLc

L'analisi dei risultati per la metrica MDLc (Figura 4.2) conferma il suo comportamento atteso nel distinguere la complessità significativa dal rumore. L'immagine di **rumore casuale** (4.11) riceve un punteggio di complessità molto basso, paragonabile a quello della copertina quasi bianca dei Beatles (4.13), dimostrando la capacità dell'algoritmo di ignorare il disordine non strutturato. Al contrario, la copertina con il **pattern geometrico** (4.12) ottiene un punteggio notevolmente alto, indicando una sensibilità della metrica ai dettagli locali e ai bordi ad alta frequenza. Le altre copertine del set, che presentano una maggiore varietà visiva, si posizionano come previsto in un range di complessità medio-alto.



(1) Rumore casuale



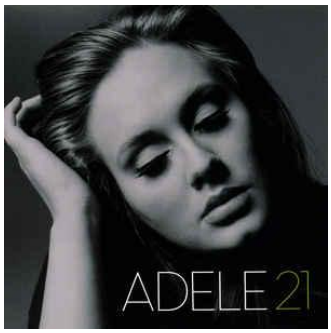
(2) Pulse Demon, Merzbow

(3) The White Album, The Beatles

(4) Bitches Brew, Miles Davis

(5) Gargoyle, Mark Lanegan Band

(6) Jackyl, Jackyl



(7) 21, Adele

(8) Frehley's Comet, Ace Frehley

(9) Blue's, Zucchero Fornacciari

Figura 4.1: Griglia di 9 immagini campione per l'analisi qualitativa delle metriche di complessità.

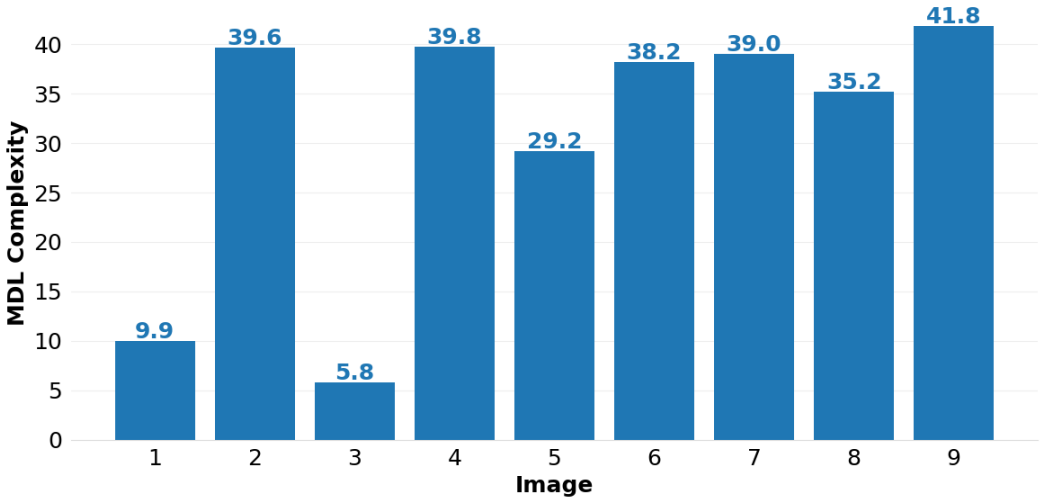


Figura 4.2: Complessità delle 9 immagini campione calcolata con la metrica MDLc. Valori più alti indicano una maggiore complessità significativa.

4.4.2 Analisi della Metrica ZIPc

I risultati per la metrica di compressione ZIPc (Figura 4.3) evidenziano un approccio differente ma con risultati nel complesso simili. Come previsto, l'immagine di **rumore casuale** (4.11) ottiene il punteggio di complessità più alto, essendo massimamente incomprimibile. Anche in questo caso, la copertina quasi bianca (4.13) viene correttamente identificata come priva di complessità, con un punteggio prossimo allo zero. È interessante notare come la copertina con il pattern geometrico (4.12) riceva un punteggio intermedio, riflettendo la sua natura parzialmente ridondante e quindi comprimibile. Le altre copertine si distribuiscono coerentemente tra questi estremi.

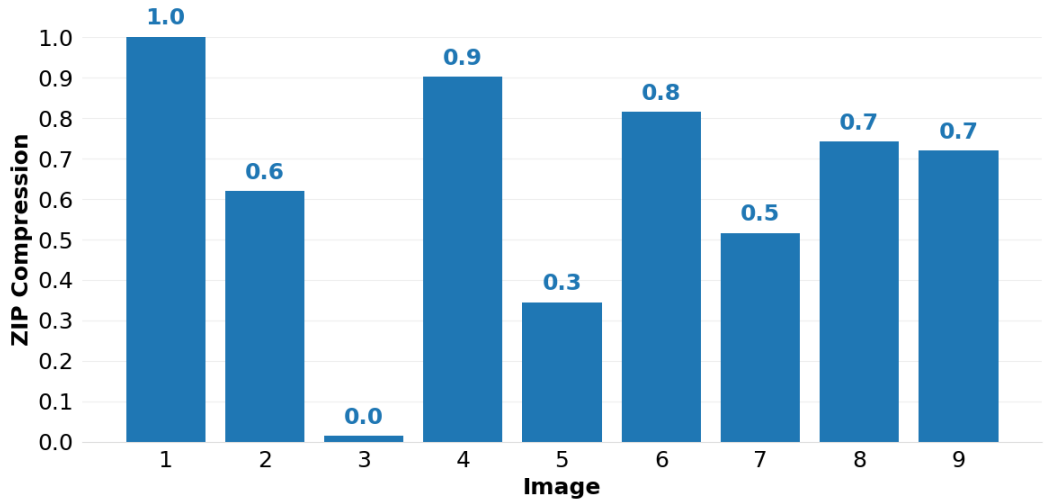


Figura 4.3: Complessità delle 9 immagini campione calcolata tramite il rapporto di compressione ZIP (ZIPc). Un valore più alto indica una minore comprimibilità e, quindi, una maggiore complessità secondo questa metrica.

4.4.3 Analisi del Piano Entropia-Complessità

Il posizionamento delle immagini nel Piano Entropia-Complessità (Figura 4.4) offre l'analisi più sfaccettata, distinguendo tra disordine e struttura. I risultati sono pienamente coerenti con la teoria:

- Il **rumore casuale** (4.11) si posiziona correttamente a massima entropia (H) e minima complessità statistica (C).
- La copertina **quasi bianca** (4.13) occupa l'estremo opposto, con valori minimi per entrambe le misure.
- La copertina con il **pattern ripetitivo** (4.12), essendo molto ripetitiva, si colloca in una regione a bassa complessità statistica ma mantiene valori alti di entropia grazie ai frequesnti cambiamenti nei pattern locali.

Un risultato particolarmente significativo riguarda la copertina dell'album *Bitches Brew* (4.14). L'opera, un dipinto dell'artista Mati Klarwein riconducibile alla corrente surrealista, si posiziona nel Piano Entropia-Complessità in una regione che corrisponde fedelmente a quella dello stile Surrealista, come analiz-

zato nel lavoro originale di Sigaki et al. Questa forte corrispondenza fornisce una solida validazione empirica della corretta implementazione della metrica.

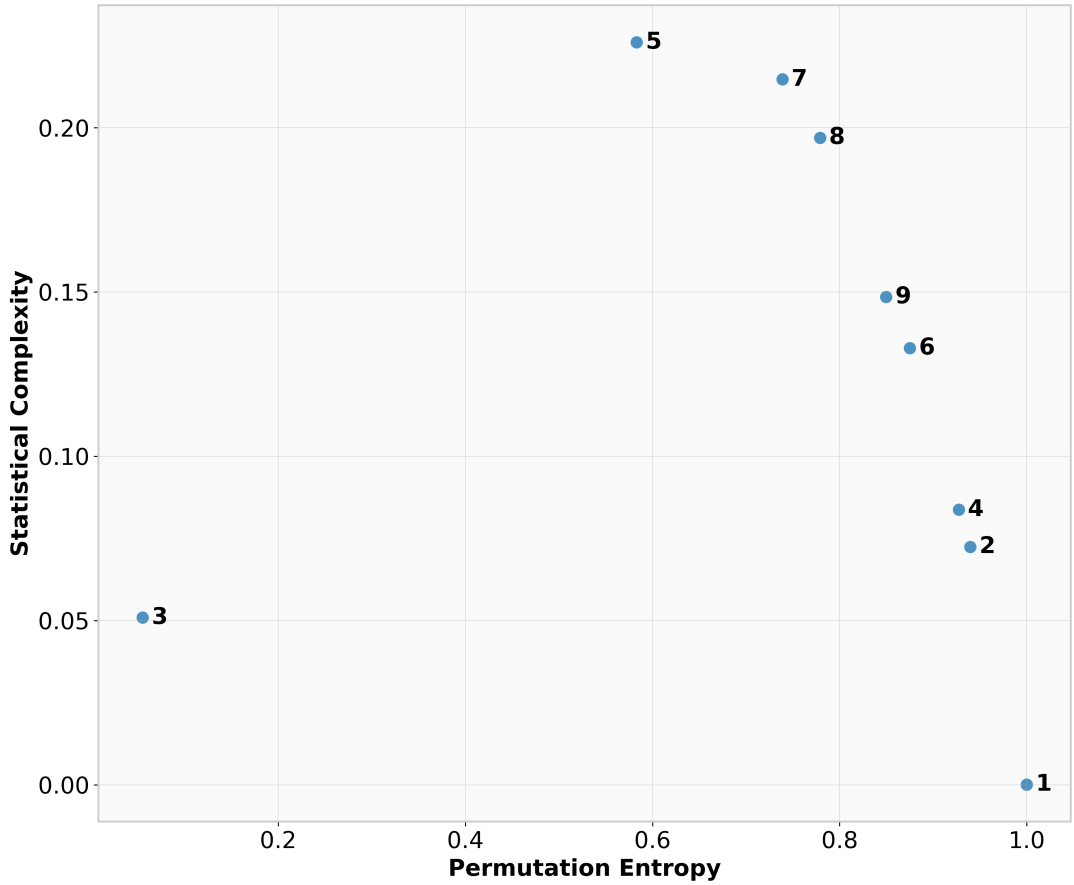


Figura 4.4: Posizionamento delle 9 immagini campione nel Piano Entropia-Complessità. Il piano distingue tra disordine casuale (alta entropia, bassa complessità) e complessità strutturata (alta complessità statistica).

4.5 Object Detection con YOLO

Per integrare le metriche di complessità percettiva con una misura di tipo semantico, è stato utilizzato un approccio basato sull'identificazione di oggetti. L'ipotesi di fondo è che un'immagine contenente un maggior numero di oggetti riconoscibili, o una maggiore varietà di essi, possa essere considerata semanticamente più complessa.

Per questo compito è stato scelto il modello **YOLOv8** ("You Only Look Once", versione 8), un sistema di object detection allo stato dell'arte sviluppato da

Ultralytics [39]. YOLO è un modello *one-stage*, noto per il suo eccellente bilanciamento tra velocità e accuratezza. È stato utilizzato il modello pre-addestrato `yolov8n.pt`, la versione più leggera e veloce, ottimizzata per l'inferenza rapida.

Il modello è addestrato sul dataset **COCO** (Common Objects in Context), il che lo rende un sistema di tipo *closed-set*: è in grado di riconoscere unicamente un set predefinito di 80 categorie (4.1) di oggetti comuni (es. persone, animali, veicoli, strumenti musicali). Da ogni immagine sono state quindi estratte le seguenti stime di complessità: il **numero totale di oggetti** rilevati e il **numero di classi di oggetti uniche**.

person	bicycle	car	motorcycle
airplane	bus	train	truck
boat	traffic light	fire hydrant	stop sign
parking meter	bench	bird	cat
dog	horse	sheep	cow
elephant	bear	zebra	giraffe
backpack	umbrella	handbag	tie
suitcase	frisbee	skis	snowboard
sports ball	kite	baseball bat	baseball glove
skateboard	surfboard	tennis racket	bottle
wine glass	cup	fork	knife
spoon	bowl	banana	apple
sandwich	orange	broccoli	carrot
hot dog	pizza	donut	cake
chair	couch	potted plant	bed
dining table	toilet	tv	laptop
mouse	remote	keyboard	cell phone
microwave	oven	toaster	sink
refrigerator	book	clock	vase
scissors	teddy bear	hair drier	toothbrush

Tabella 4.1: Elenco delle 80 classi di oggetti riconosciute da YOLOv8.

4.5.1 Implementazione Pratica

L'analisi è stata implementata in Python utilizzando la libreria ufficiale `ultralytics`. È stato impiegato il modello pre-addestrato `yolov8n.pt`, impostando una **soglia di confidenza (confidence threshold) di 0.25** per considerare valida una rilevazione.

Per ogni immagine del dataset, lo script ha registrato:

- Il numero totale di oggetti identificati.
- L'elenco e il conteggio delle classi di oggetti uniche (es. {'person': 2, 'guitar': 1}).
- La percentuale totale dell'area dell'immagine coperta dai riquadri di delimitazione (bounding box) degli oggetti rilevati.

I risultati sono stati poi uniti al dataset principale. Data la natura computazionalmente intensiva dell'analisi su decine di migliaia di immagini, il processo è stato ottimizzato tramite parallelismo, come descritto nella Sezione 5.

Capitolo 5

Strumenti Utilizzati

Questo capitolo descrive gli strumenti software e le tecniche computazionali impiegate per implementare le metodologie di analisi, etichettatura e download dei dati descritte nei capitoli precedenti. L'obiettivo è fornire un quadro tecnico completo a supporto della piena riproducibilità della ricerca.

Tutto il codice sviluppato per questa tesi è open source e disponibile nella seguente repository GitHub: Album Cover Complexity

5.1 Python e librerie

L'intero progetto è stato sviluppato in linguaggio **Python** (versione 3.9+). La scelta è stata motivata dalla sua versatilità e dal vasto ecosistema di librerie per l'analisi scientifica, il machine learning e la manipolazione dei dati. Le librerie chiave impiegate includono:

- **Pandas:** Per la gestione e l'analisi dei dati tabulari. È stata utilizzata per leggere, modificare e salvare tutti i dataset in formato CSV.
- **NumPy:** Per le operazioni numeriche vettorializzate e la gestione efficiente di matrici, fondamentale per l'elaborazione delle immagini e per l'implementazione degli algoritmi di complessità.
- **Pillow (PIL):** Per l'apertura, la manipolazione (es. ridimensionamento) e la conversione delle immagini tra diversi formati.

- **Matplotlib e Seaborn:** Per la generazione di tutti i grafici e le visualizzazioni presenti in questa tesi.
- **Tqdm:** Per la creazione di barre di avanzamento, utili a monitorare l'esecuzione di processi computazionalmente intensivi.

5.2 Esecuzione in Batch e Checkpointing

Diverse operazioni, in particolare l'interazione con servizi esterni (come le API Google AI), sono state strutturate in **batch** (lotti). Questo approccio consiste nel raggruppare un certo numero di richieste o di dati (ad esempio, 100 album alla volta) e processarli insieme, permettendo di ottimizzare le chiamate di rete e di gestire la memoria in modo più efficiente.

Oltre a questi vantaggi, l'approccio a batch è stato abbinato a una tecnica di **salvataggio progressivo** per garantire la robustezza dei processi di lunga durata. Durante le elaborazioni più estese, come il calcolo delle metriche di complessità, i risultati intermedi venivano salvati su disco in file CSV a intervalli regolari (es. ogni 5000 righe). Questa strategia ha reso il sistema resiliente agli errori: in caso di interruzione inaspettata, è stato possibile riprendere l'elaborazione dall'ultimo checkpoint, evitando la perdita di ore di lavoro e di dati già acquisiti.

5.3 Ottimizzazione tramite Parallelismo

Diverse fasi del progetto, in particolare l'analisi delle immagini per il calcolo delle metriche di complessità (MDLc, Entropia-Complessità) e l'esecuzione del modello YOLOv8, sono computazionalmente onerose. Per ridurre drasticamente i tempi di esecuzione sull'intero dataset, è stato implementato il **parallelismo dei processi**. Utilizzando i moduli `multiprocessing` e `concurrent.futures` di Python, è stato possibile distribuire il carico di lavoro su tutti i core della CPU disponibili, eseguendo l'analisi di più immagini in contemporanea e abbattendo i tempi di elaborazione in maniera considerevole.

Capitolo 6

Risultati

In questo capitolo vengono presentati e discussi i risultati ottenuti dall'analisi della complessità visiva delle copertine degli album musicali. L'indagine è strutturata in due sezioni: la prima esplora le differenze di complessità tra i generi musicali, mentre la seconda ne analizza l'evoluzione temporale.

6.1 Analisi della Complessità per Genere Musicale

L'obiettivo primario di questa sezione è determinare se la complessità visiva possa essere considerata una caratteristica discriminante per i diversi generi musicali.

L'analisi della distribuzione dei punteggi di complessità, basata sulla metrica MDLc (Figura 6.1), rivela un'omogeneità tra la maggior parte dei generi. Come si può vedere dal grafico, le distribuzioni dei punteggi risultano affini, con valori medi che si attestano su livelli tra loro molto vicini. Questa osservazione potrebbe suggerire l'esistenza di una "norma" stilistica trasversale a cui la maggior parte delle copertine tende a conformarsi, indipendentemente dal genere.

Tuttavia, emergono due eccezioni significative: **Metal** e **Hip Hop**. Entrambi i generi presentano distribuzioni dei valori di complessità traslate verso l'alto, indicando una maggiore densità di informazione visiva nelle loro copertine. Tale risultato è ulteriormente corroborato dall'analisi basata sulla metrica ZIPc (Fi-

gura 6.2), nella quale il genere Metal si conferma come quello con la complessità mediamente più elevata.

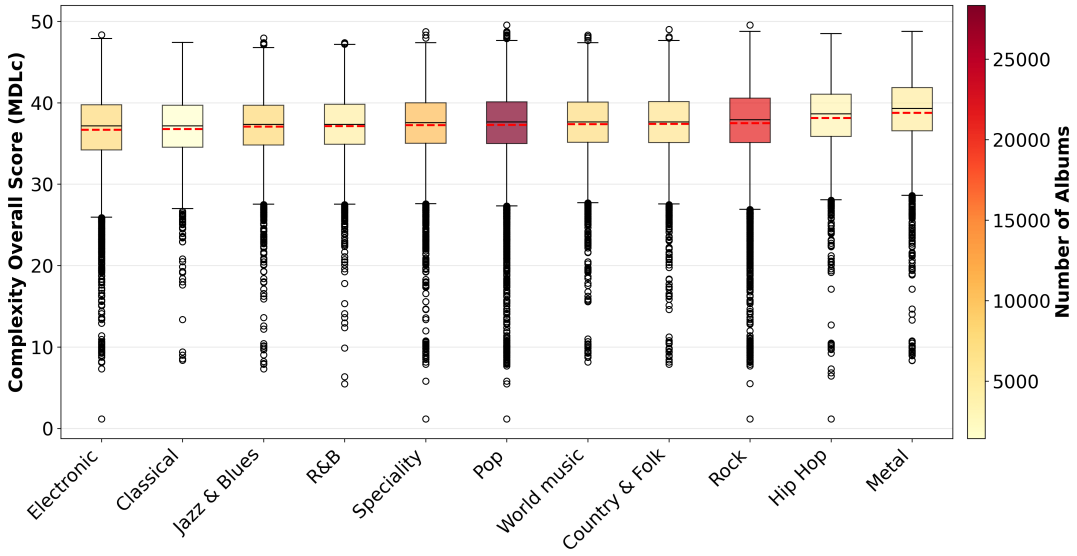


Figura 6.1: Box plot della distribuzione del punteggio di complessità MDLc per ciascuno degli 11 generi. La linea nera indica la mediana, la linea rossa tratteggiata la media. La colorazione è proporzionale al numero di album per genere.

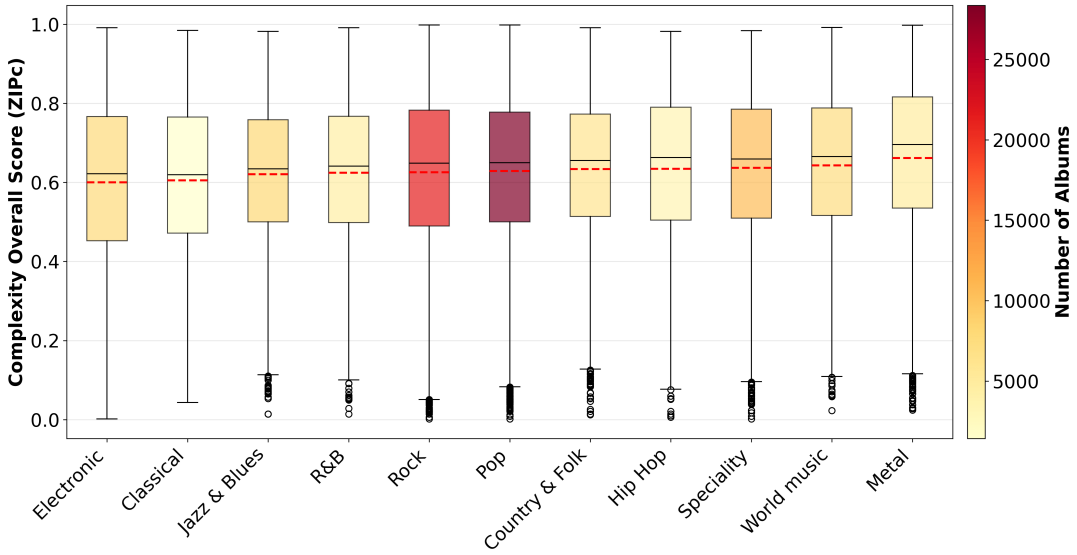


Figura 6.2: Box plot della distribuzione del rapporto di compressione (ZIPc) per genere.La linea nera indica la mediana, la linea rossa tratteggiata la media. La colorazione è proporzionale al numero di album per genere.

Un'altra osservazione rilevante riguarda la variabilità interna di alcuni generi. **Electronic** e **Rock** mostrano un range interquartile (l'ampiezza della scatola)

e una dispersione dei dati (i "baffi") più ampi rispetto agli altri. Questo suggerisce una maggiore eterogeneità stilistica interna. Una possibile spiegazione per questo potrebbe essere la coesistenza di diverse correnti artistiche all'interno dello stesso genere.

Infine, un fenomeno trasversale è la massiccia presenza di *outlier* nella parte inferiore della distribuzione per tutti i generi. Tale osservazione evidenzia come la realizzazione di copertine a bassa complessità costituisca una scelta stilistica ampiamente diffusa, che trascende i singoli generi musicali.

Un'ulteriore prospettiva è offerta dal posizionamento dei generi nel piano entropia-complessità (Figura 6.3). Il grafico mostra che la maggior parte dei generi si aggrega in un *cluster* centrale, confermando l'idea di una norma condivisa. Le eccezioni già osservate vengono qui ulteriormente caratterizzate:

- **Metal** e **Hip Hop** si confermano come *outlier*, posizionandosi in una regione a più alta Entropia di Permutazione (H). Questo suggerisce che la loro elevata complessità è associata a una maggiore casualità e imprevedibilità nella struttura visiva.

- **Electronic** si colloca in una posizione quasi opposta, con un valore di entropia inferiore ma una Complessità Statistica (C) elevata, un profilo tipico di sistemi con strutture complesse ma ordinate riconducibili ad un estetica minimalista.

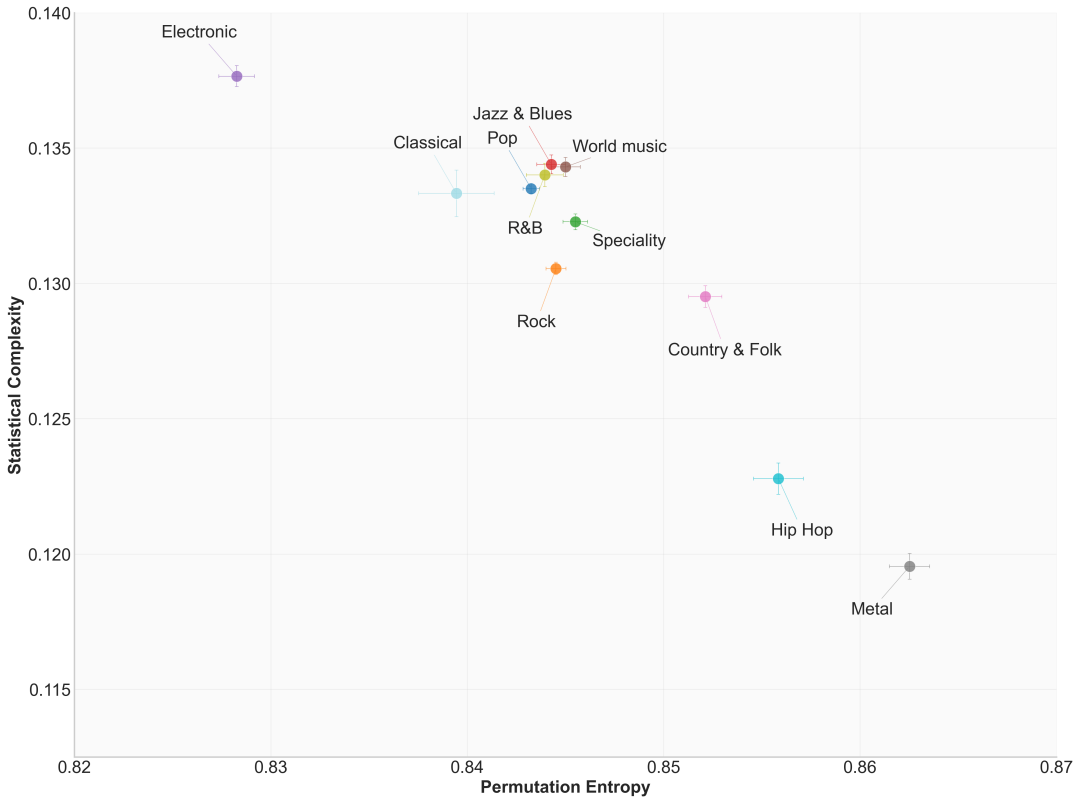


Figura 6.3: Posizionamento dei valori medi di Entropia di Permutazione (asse x) e Complessità Statistica (asse y) per ciascun genere. Le barre di errore indicano l'errore standard della media (SEM).

Per integrare le metriche di complessità percettiva con una misura basata sul contenuto figurativo, è stato utilizzato un approccio fondato sull'identificazione di oggetti utilizzando il modello **YOLOv8**. L'ipotesi di fondo è che un'immagine contenente un maggior numero di elementi riconoscibili, o una maggiore varietà di essi, possa essere considerata più ricca dal punto di vista contenutistico.

Dall'analisi è emerso che la classe di gran lunga più rilevata è stata **person**, suggerendo che la figura umana è l'elemento più comune sulle copertine degli album. La lista completa delle occorrenze per ogni classe è riportata in dettaglio nella Tabella 6.1.

Passando invece all'analisi della distribuzione della presenza di oggetti tra i generi musicali, illustrata nella Figura 6.4, si osserva che la maggioranza delle copertine, per quasi tutti i generi, contiene almeno un oggetto riconoscibile Generi come **R&B** (81.8% di album con oggetti), **Country & Folk** (79.0%) e **Hip Hop** (76.3%) presentano le più alte percentuali di copertine con elementi

Classe	Occ.	Classe	Occ.	Classe	Occ.	Classe	Occ.
person	55746	sports ball	324	boat	160	couch	75
tie	2535	frisbee	323	airplane	153	sheep	67
clock	1881	kite	289	toothbrush	133	elephant	61
bird	842	suitcase	275	bicycle	132	hot dog	57
car	829	bed	274	handbag	130	bear	50
dog	770	truck	268	dining table	128	oven	49
train	672	traffic light	264	giraffe	122	microwave	46
chair	646	donut	259	bowl	112	parking meter	43
vase	617	skateboard	248	apple	109	knife	42
bottle	585	motorcycle	218	remote	108	carrot	42
cat	580	bench	215	orange	106	broccoli	41
book	552	banana	207	toilet	104	skis	36
scissors	537	stop sign	207	baseball bat	101	mouse	35
horse	531	bus	205	pizza	100	backpack	33
tv	524	tennis racket	193	baseball glove	98	sink	21
umbrella	489	wine glass	186	fire hydrant	95	fork	19
cup	397	potted plant	182	refrigerator	92	sandwich	17
cell phone	371	keyboard	179	zebra	91	spoon	10
teddy bear	345	cow	174	laptop	91	toaster	1
cake	331	surfboard	161	snowboard	81		

Tabella 6.1: Occorrenze per le 79 classi di oggetti rilevate nel dataset, ordinate in modo decrescente. In totale, sono stati identificati 77.422 oggetti.

rilevabili. Questa alta incidenza è probabilmente legata alla forte tradizione di questi generi di mettere in primo piano l'artista sulla copertina, come suggerito dalla predominanza della classe "person" nei rilevamenti totali.

Al contrario, i generi **Metal**, **Electronic** e **Rock** sono quelli con la più bassa percentuale di album contenenti oggetti riconoscibili (rispettivamente 55.2%, 63.7% e 66.5%). Le peculiarità di **Metal** ed **Electronic**, in particolare, saranno oggetto di un'analisi più approfondita nella Sezione 6.2, dove i dati sul contenuto figurativo offriranno spunti significativi per interpretare i risultati delle metriche di complessità percettiva.

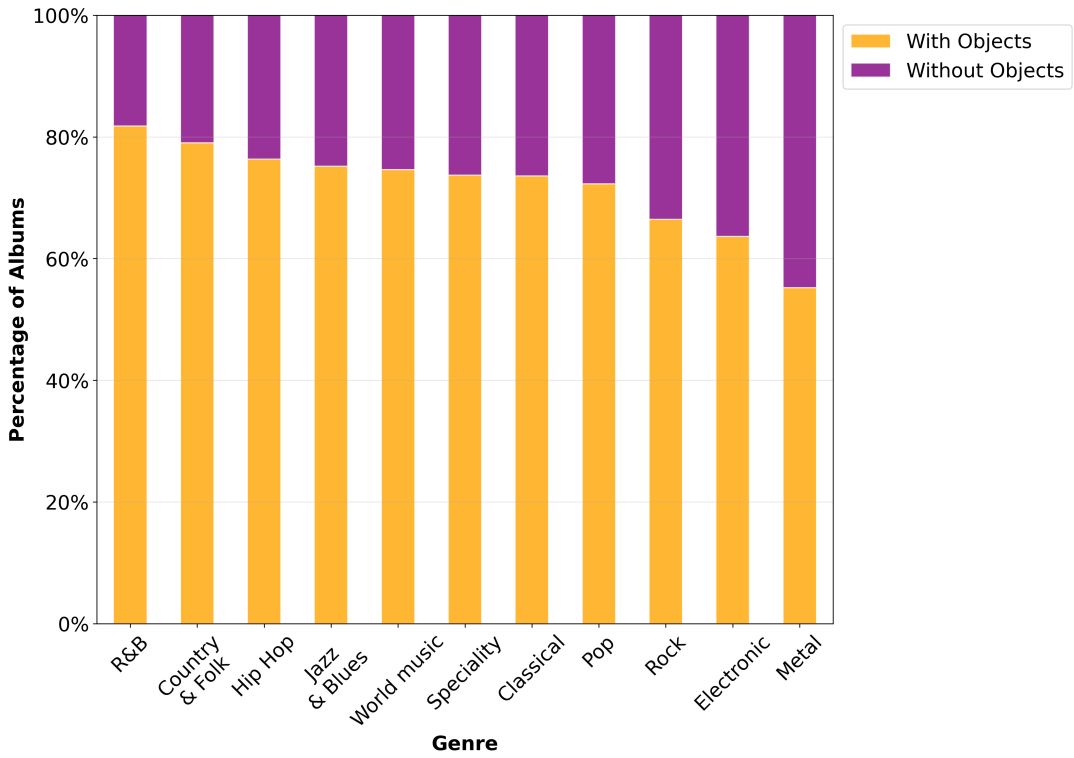


Figura 6.4: Percentuale di album con oggetti (in arancione, parte inferiore) e senza oggetti (in viola, parte superiore) rilevabili da YOLOv8, suddivisi per genere musicale.

6.2 Evoluzione della Complessità nel Tempo

Questa sezione analizza come la complessità visiva delle copertine sia cambiata nel corso dei decenni.

Dall'analisi dell'andamento temporale (Figura 6.5 e 6.6), si nota che i generi **Metal** e **Hip Hop** mantengono costantemente i livelli di complessità più elevati, indicando la persistenza di canoni visivi ben definiti. Un'evoluzione interessante riguarda l'Hip Hop: negli ultimi 15 anni, si osserva un marcato declino della sua complessità media, che tende ad allinearsi con quella degli altri generi. Questo fenomeno potrebbe essere interpretato come una conseguenza della sua crescente popolarità e della conseguente assimilazione da parte del *mainstream*.

Nonostante le specificità di ogni genere, è possibile identificare un andamento generale comune, articolato in quattro macro-periodi:

- **Fino al 1996:** Un periodo caratterizzato da una **crescita** generalizzata della complessità media (confermato da entrambe le metriche).

- **1996-2003:** Si osserva un'inversione di tendenza, con un **calo graduale** della complessità. La metrica ZIPc evidenzia una diminuzione ancora più accentuata.

- **2003-2009:** La complessità tende a **stabilizzarsi**, con lievi oscillazioni verso l'alto.

- **Dal 2009 al 2025:** Si registra un netto e diffuso **calo della complessità** che interessa la quasi totalità dei generi, con la sola eccezione del Jazz & Blues che mantiene una lieve crescita. Questa tendenza generale è sostanzialmente confermata da entrambe le metriche, sebbene l'analisi basata su ZIPc rivela una parziale controtendenza per i generi Electronic e Metal, mostrando per essi una risalita negli anni più recenti.

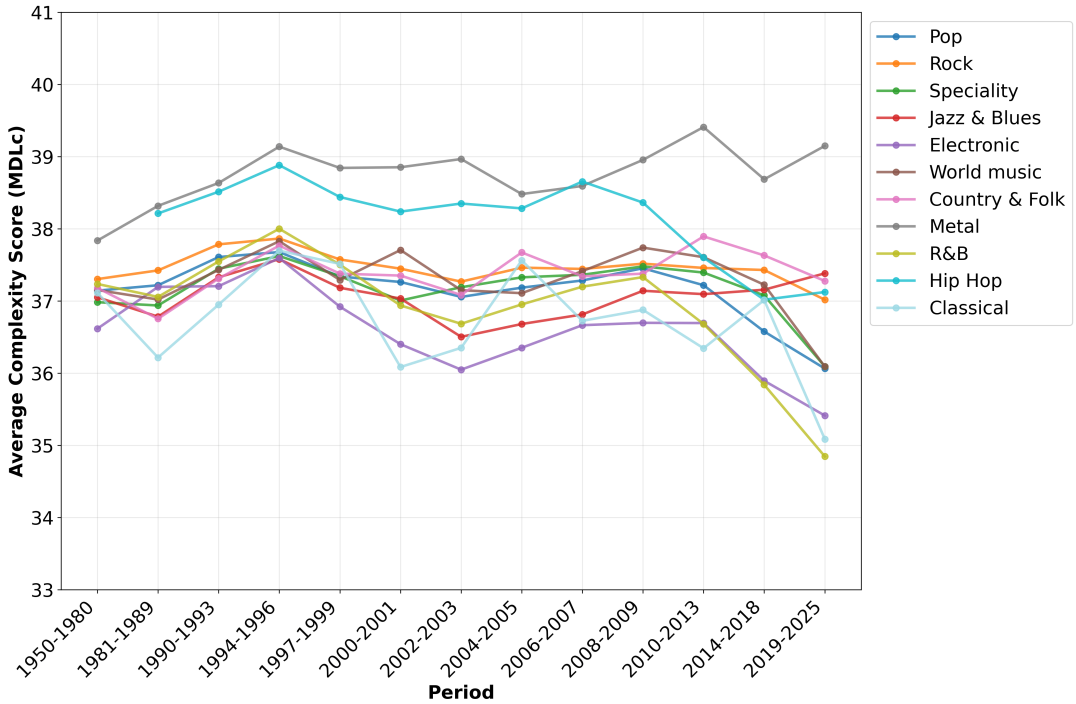


Figura 6.5: Andamento della complessità media (MDLc) nel tempo, suddivisa per genere. L'asse temporale è partizionato in periodi dinamici, ciascuno contenente almeno 3000 album totali (ad eccezione dell'ultimo, con 2269). Un genere viene mostrato in un dato periodo solo se sono presenti almeno 50 album.

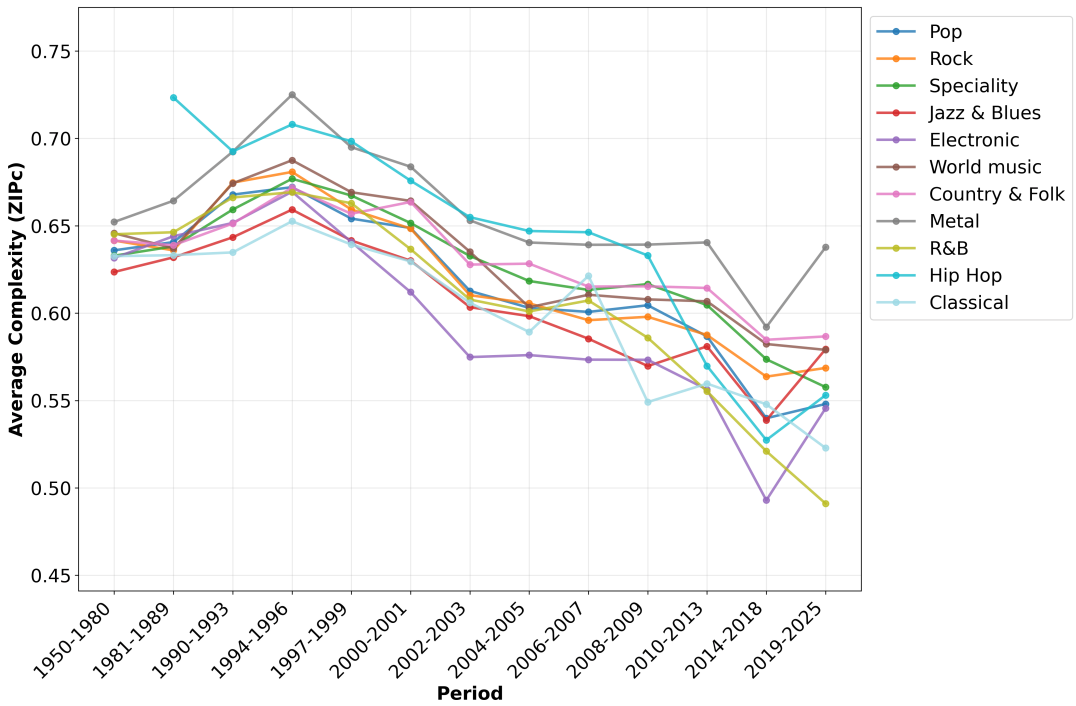


Figura 6.6: Andamento della complessità media (ZIPc) nel tempo, suddivisa per genere. L'asse temporale è partizionato in periodi dinamici, ciascuno contenente almeno 3000 album totali (ad eccezione dell'ultimo, con 2269). Un genere viene mostrato in un dato periodo solo se sono presenti almeno 50 album.

L'evoluzione temporale aggregata, visualizzata nel piano entropia-complessità (Figura 6.7), corrobora questa suddivisione:

- **Fino al 1996:** Si osserva un aumento dell'entropia a fronte di una diminuzione della complessità statistica, indicando una transizione verso copertine visivamente più disordinate.
- **1996-2003:** Una netta diminuzione dell'entropia, accompagnata da un aumento della complessità statistica, suggerisce uno spostamento verso un'estetica più strutturata e meno casuale.
- **2003-2009:** I valori rimangono relativamente stabili, indicando un consolidamento degli stili visivi.
- **Dal 2009 al 2025:** Si osserva un marcato calo della complessità statistica, a fronte di un'entropia che rimane sostanzialmente invariata.

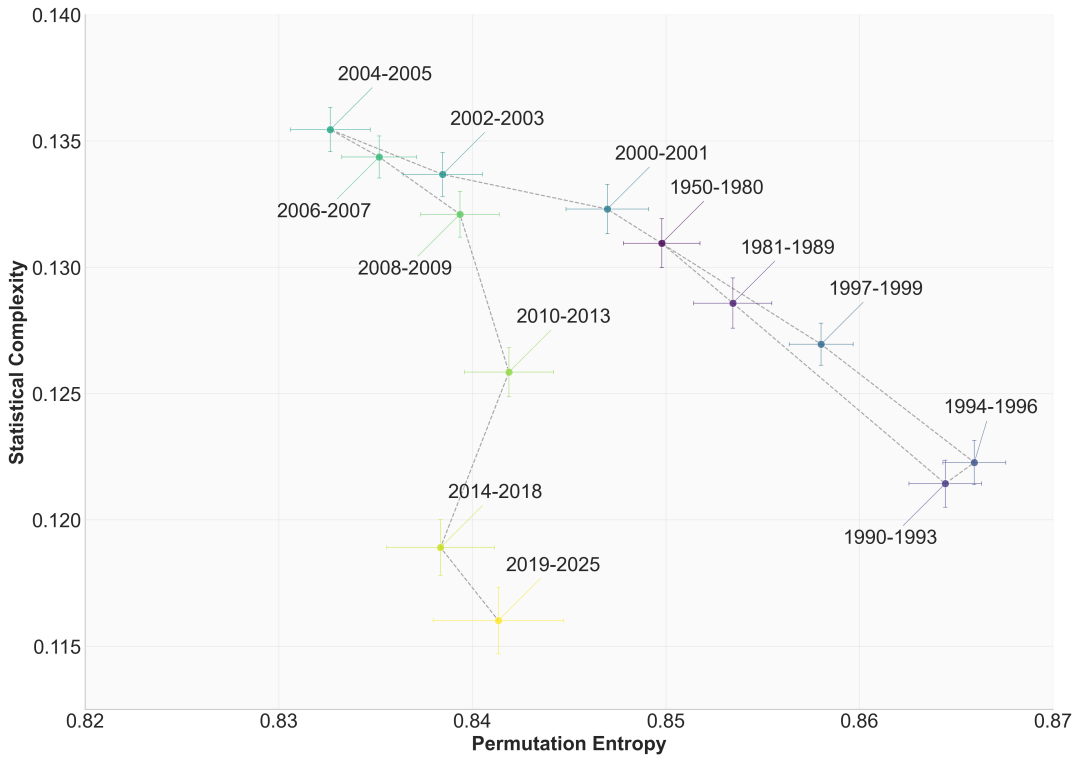


Figura 6.7: Evoluzione dei valori medi di Entropia di Permutazione e Complessità Statistica nel tempo. I punti rappresentano periodi temporali dinamici, ciascuno contenente un minimo di 3000 album ad eccezione dell'ultimo periodo che ne comprende 2269. La linea tratteggiata collega i punti in ordine cronologico, mostrando la traiettoria evolutiva aggregata delle copertine.

Spostando l'attenzione sui risultati prodotti dall'identificazione degli oggetti, l'analisi tramite YOLO si rivela uno strumento efficace per **interpretare** la complessità, poiché permette di indagare le cause dei livelli di complessità emersi dalle metriche percettive.

L'analisi dell'evoluzione temporale del numero medio di oggetti per album (Figura 6.8) offre un primo spunto cruciale. A differenza delle metriche precedenti (Figure 6.5 e 6.6), non emergono qui i chiari macro-periodi storici. Tuttavia, si conferma una tendenza generale dei generi a convergere verso una "norma" condivisa, fluttuando tra 1.5 e 2 oggetti medi per copertina.

È proprio analizzando le eccezioni a questa norma che l'approccio basato sul contenuto figurativo si rivela più utile.

Il genere **Metal** rappresenta il caso di studio più significativo. Pur essendo risultato tra i generi con la più alta complessità percettiva, esso mostra costan-

temente la media più bassa di oggetti rilevabili. I risultati suggeriscono che le copertine Metal possiedono un'elevata complessità percettiva, la cui natura, secondo la Figura 6.3, appare legata a immagini ricche di texture, dettagli minuti e disordine.

Questa idea è confermata dai risultati di YOLO: l'estetica del genere sembra composta da elementi che, pur essendo visivamente complessi, non rientrano nelle 80 classi di oggetti comuni riconosciute dal modello. Ciò è probabilmente dovuto alla predilezione per illustrazioni fantasy, loghi illeggibili, arte astratta e, in generale, composizioni prive di una struttura figurativa ben definita. Il basso numero di oggetti, quindi, non indica semplicità, ma conferma un'estetica che fonda la sua complessità sull'astrattismo e sul caos visivo piuttosto che sulla rappresentazione strutturata di oggetti reali.

Un'altra visione interessante emerge per il genere **Electronic**, dove il numero medio di oggetti rilevati è nuovamente tra i più bassi. A differenza del Metal, però, questo dato confrontato con i risultati precedenti non sembra suggerire un caos visivo, ma potrebbe piuttosto confermare la nota **natura minimalista** di molte sue copertine. Questa interpretazione è coerente con i bassi valori di complessità già evidenziati nelle Figure 6.3, 6.1 e 6.5, suggerendo la presenza di artwork con pochi oggetti o basato su forme geometriche pulite, gradienti di colore e pura tipografia.

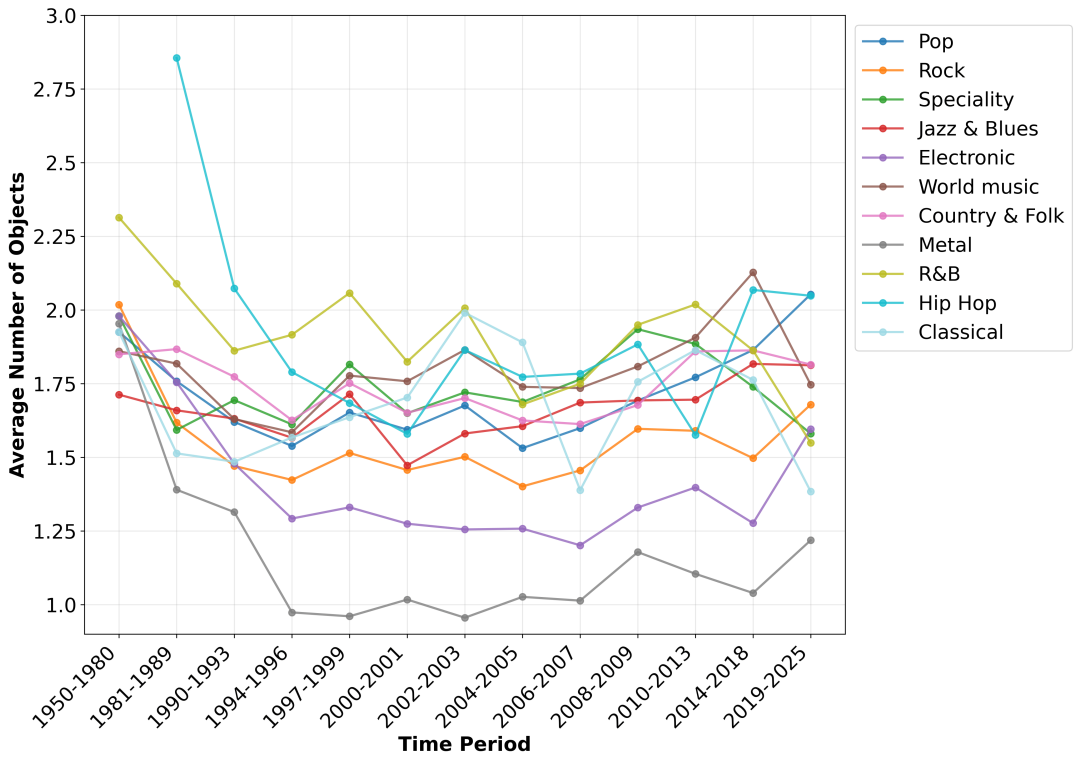


Figura 6.8: Evoluzione nel tempo del numero medio di oggetti per album riconosciuti da YOLOv8, suddivisi per genere. L'asse temporale è partizionato in periodi dinamici, ciascuno contenente almeno 3000 album totali (ad eccezione dell'ultimo, con 2269). Un genere viene mostrato in un dato periodo solo se sono presenti almeno 50 album.

Infine, un'osservazione di particolare interesse emerge dall'analisi della distribuzione dei punteggi di complessità all'interno dei diversi periodi (Figura 6.9 e 6.10). Nonostante le medie mostrino le tendenze descritte, la **variabilità** dei punteggi rivela a sua volta una dinamica di interesse. Si nota un aumento progressivo della dispersione (visibile dall'ampliarsi delle scatole), particolarmente accentuato negli ultimi decenni. Questo aumento della dispersione potrebbe essere interpretato come un indicatore di una crescente differenziazione stilistica, suggerendo che, specialmente negli ultimi anni, le estetiche delle copertine tendano a evolvere in direzioni sempre più eterogenee.

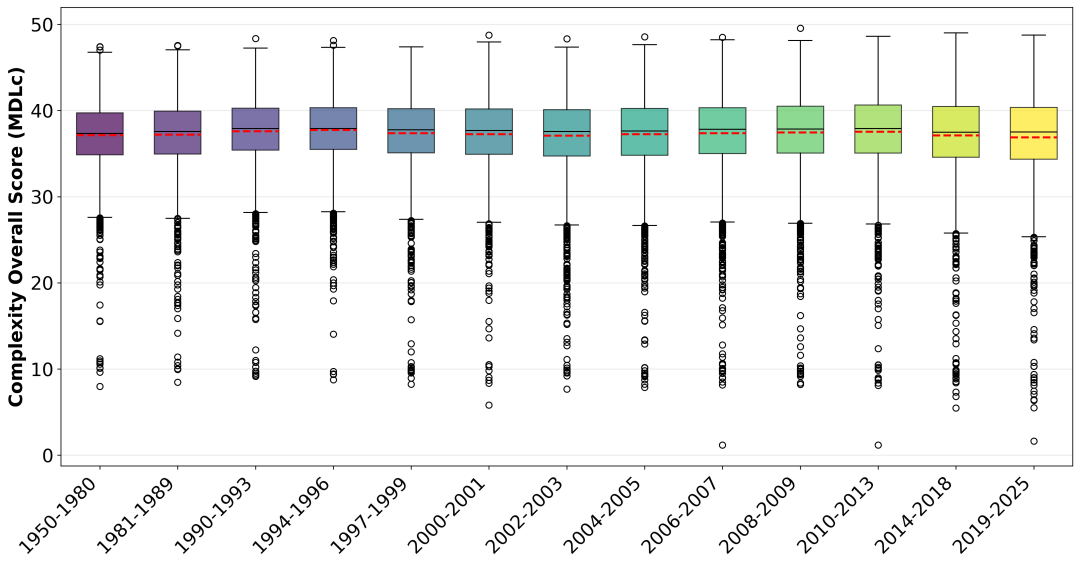


Figura 6.9: Box plot della distribuzione della complessità MDLc per periodo temporale. I dati sono raggruppati in periodi dinamici contenenti almeno 3000 album (eccetto l'ultimo che ne contiene 2269).

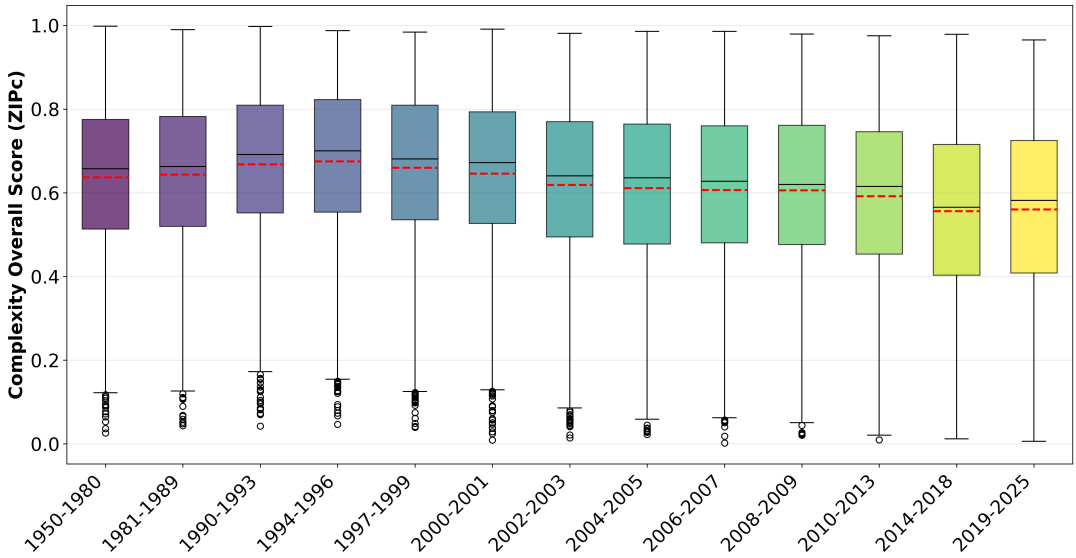


Figura 6.10: Box plot della distribuzione del rapporto di compressione (ZIPc) per periodo temporale, con i dati raggruppati in periodi dinamici contenenti almeno 3000 album (eccetto l'ultimo che ne contiene 2269).

Capitolo 7

Conclusioni

Il presente lavoro di tesi si è proposto di condurre un'analisi quantitativa su larga scala dell'evoluzione della complessità visiva nelle copertine degli album musicali. L'obiettivo era duplice: investigare le differenze tra i generi musicali e, soprattutto, tracciare le traiettorie evolutive di questi artefatti culturali nel tempo.

Per raggiungere tale scopo, è stato costruito un dataset di oltre 46.000 album, frutto della fusione di tre diverse fonti dati e arricchito tramite l'impiego di Modelli Linguistici di Grandi Dimensioni. L'analisi è stata fondata su un approccio multi-metrico, volto a catturare diverse sfaccettature della nozione di "complessità". Sono state impiegate metriche di complessità percettiva, come la complessità significativa basata sul principio della Minima Lunghezza di Descrizione (*MDLc*), il rapporto di compressione (*ZIPc*) e il Piano Entropia-Complessità, affiancate da un'analisi del contenuto figurativo basata sul riconoscimento di oggetti tramite *YOLOv8*.

I risultati hanno rivelato intuizioni significative. In particolare, l'analisi per genere ha mostrato l'esistenza di una "norma" stilistica trasversale a cui la maggior parte delle copertine tende a conformarsi. Questa norma è quantificabile attraverso diverse metriche: la complessità MDLc si attesta tra 35 e 40 e il rapporto di compressione tra 0.6 e 0.66, mentre l'entropia di permutazione è di circa 0.84, con una complessità statistica compresa tra 0.130 e 0.135. Tuttavia, generi come *Metal*, *Hip Hop* ed *Electronic* sono emersi come outlier, mostrando

però caratteristiche opposte. Da un lato, *Metal* e *Hip Hop* sono caratterizzati da una complessità visiva mediamente più elevata. L'analisi del Piano Entropia-Complessità arricchisce questa osservazione, associando la loro complessità a un più alto grado di disordine (entropia). Dall'altro lato, l'*Electronic* si distingue per una complessità inferiore alla media che, secondo la stessa analisi, deriverebbe da strutture visive più ordinate e minimaliste.

L'analisi del contenuto figurativo, infine, è stata cruciale per interpretare il caso specifico del *Metal*, confermando che la sua elevata complessità percettiva è probabilmente dovuta a un'estetica astratta e caotica.

L'analisi temporale ha costituito il nucleo di questa ricerca, portando alla luce un andamento generale condiviso dalla maggior parte dei generi e articolato in quattro macro-periodi. Il contributo più rilevante è stata l'identificazione di un netto e diffuso calo della complessità media a partire dal 2009. Questa tendenza è particolarmente evidente nella complessità statistica, che ha registrato una diminuzione di circa il 12% negli ultimi 15 anni. Questa tendenza sembra avvalorare l'ipotesi iniziale secondo cui i cambiamenti nei paradigmi di fruizione musicale, in particolare la transizione verso le piattaforme digitali e la visualizzazione su schermi di piccole dimensioni, possano aver esercitato una pressione selettiva verso un'estetica visivamente più semplice e d'impatto immediato.

L'analisi ha rivelato anche una dinamica più sottile. Nello stesso periodo in cui la complessità media diminuiva, si è infatti osservato un aumento della variabilità stilistica, concentrato in particolare negli ultimi 15 anni. Questa crescente polarizzazione è quantificabile, ad esempio, in un aumento di circa l'11% nella dispersione dell'IQR, un valore riscontrato sia per la metrica MDLc sia per quella basata sul rapporto di compressione. Ciò suggerisce che, piuttosto che una semplice omogeneizzazione, l'ecosistema digitale potrebbe aver favorito una polarizzazione delle estetiche: da un lato copertine minimaliste e dirette, dall'altro una continua esplorazione di nicchie visive complesse.

Questo studio, pur con i limiti legati alla costruzione del dataset, getta le basi per interessanti direzioni di ricerca future. Sarebbe innanzitutto fondamentale ampliare il dataset con un numero maggiore di album degli anni più recenti per validare le tendenze osservate e, in particolare, per investigare il potenziale

impatto dei nuovi strumenti di Intelligenza Artificiale Generativa sul design delle copertine. Un'altra direzione promettente riguarda l'integrazione di dati sul successo commerciale e sulla popolarità, per verificare se esista una correlazione tra la complessità visiva e la performance di mercato di un album. Infine, sarebbe di grande interesse correlare la complessità visiva della copertina con la complessità musicale e lirica del contenuto sonoro, per rispondere alla domanda: un'estetica semplice corrisponde a una musica più semplice?

In conclusione, l'analisi condotta in questa tesi ha quantificato l'evoluzione della complessità visiva delle copertine musicali su larga scala. I risultati ottenuti suggeriscono che l'estetica di questi artefatti non sia casuale. Piuttosto, tale evoluzione sembra riflettere una risposta misurabile alle trasformazioni dell'ecosistema che le circonda, ponendosi in continuità con le tendenze alla semplificazione già documentate sia per la musica stessa [9] sia per i testi delle canzoni [8, 23].

Bibliografia

[1] Leonard Bernstein. *The Unanswered Question: Six Talks at Harvard*. Charles Eliot Norton Lectures. Cambridge, MA: Harvard University Press, 1976.

[2] Graham F. Welch et al. «Editorial: The Impact of Music on Human Development and Well-Being». In: *Frontiers in Psychology* 11 (2020), p. 1246. DOI: 10.3389/fpsyg.2020.01246.

[3] Ian Cross. «Music, Cognition, Culture, and Evolution». In: *Annals of the New York Academy of Sciences* 930.1 (2001), pp. 28–42.

[4] Samuel A. Mehr et al. «Universality and Diversity in Human Song». In: *Science* 366.6468 (2019), eaax0868. DOI: 10.1126/science.aax0868.

[5] Steven Heller. *Alex Steinweiss, Inventor of the Modern Album Cover, Dies at 94*. The New York Times. URL: https://www.nytimes.com/2011/07/20/arts/music/alex-steinweiss-inventor-of-the-album-cover-dies-at-94.html. Lug. 2011.

[6] Steve Jones e MARTIN SORGER. «Covering music: A brief history and analysis of album cover design». In: *Journal of Popular Music Studies* 11 (ago. 2006), pp. 68–102. DOI: 10.1111/j.1533-1598.1999.tb00004.x.

[7] Mikkel Vad. «The Album Cover». In: *Journal of Popular Music Studies* 33.3 (2021), pp. 11–15. DOI: 10.1525/jpms.2021.33.3.11.

[8] Emilia Parada-Cabaleiro et al. «Song Lyrics Have Become Simpler and More Repetitive Over the Last Five Decades». In: *Scientific Reports* 14.1 (2024). Erratum in: Sci Rep. 2024 May 22;14(1):11712, p. 5531. DOI: 10.1038/s41598-024-55742-x.

[9] Niccolò Di Marco et al. «Decoding Musical Evolution Through Network Science». In: *arXiv preprint arXiv:2501.07557* (2025). arXiv: 2501.07557 [[cs.SD](http://cs.sd/)].

[10] Don Donderi. «Visual Complexity: A Review». In: *Psychological bulletin* 132 (gen. 2006), pp. 73–97. DOI: 10.1037/0033-2909.132.1.73.

[11] Ming Li e Paul Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer-Verlag, 1997.

[12] Ruth Rosenholtz, Yuanzhen Li e Lisa Nakano. «Measuring visual clutter». In: *Journal of Vision* 7.2 (2007), p. 17. DOI: 10.1167/7.2.17.

[13] Manuela M. Marin e Helmut Leder. «Examining Complexity Across Domains: Relating Subjective and Objective Measures of Affective Environmental Scenes, Paintings and Music». In: *PLoS ONE* 8.8 (2013), e72412. DOI: 10.1371/journal.pone.0072412.

[14] Tariq Khan, Syed Naqvi e Erik Meijering. «Leveraging image complexity in macro-level neural network design for medical image segmentation». In: *Scientific Reports* 12 (dic. 2022). DOI: 10.1038/s41598-022-26482-7.

[15] Hongyi Yang e Han Yang. «Evolution of Entropy in Art Painting Based on the Wavelet Transform». In: *Entropy* 23.7 (2021), p. 883. DOI: 10.3390/e23070883.

[16] Christoph Redies et al. «PHOG-Derived Aesthetic Measures Applied to Color Photographs of Artworks, Natural Scenes and Objects». In: vol. 7583. Ott. 2012. ISBN: 978-3-642-33862-5. DOI: 10.1007/978-3-642-33863-2_54.

[17] Nina Lam et al. «An Evaluation of Fractal Methods for Characterizing Image Complexity». In: *Cartography and Geographic Information Science - CARTOGR GEOGR INF SCI* 29 (gen. 2002), pp. 25–35. DOI: 10.1559/152304002782064600.

[18] Louis Mahon e Thomas Lukasiewicz. «Minimum description length clustering to measure meaningful image complexity». In: *Pattern Recognition* 145 (2024), p. 109889. ISSN: 0031-3203. DOI: https://doi.org/10.1016/j.patcog.2023.109889. URL: https://www.sciencedirect.com/science/article/pii/S0031320323005873.

[19] Higor Y. D. Sigaki, Matjaž Perc e Haroldo V. Ribeiro. «History of art paintings through the lens of entropy and complexity». In: *Proceedings of the National Academy of Sciences of the United States of America* 115.37 (2018), E8585–E8594. DOI: 10.1073/pnas.1800083115.

[20] Christoph Adami, Charles Ofria e Travis C. Collier. «Evolution of Biological Complexity». In: *Proceedings of the National Academy of Sciences of the United States of America* 97.9 (2000), pp. 4463–4468. DOI: 10.1073/pnas.97.9.4463.

[21] Michael Batty. «Complexity in City Systems: Understanding, Evolution, and Design». In: *Batty, M. (2007) Complexity in city systems: understanding, evolution, and design. Working paper. CASA Working Papers (117). Centre for Advanced Spatial Analysis (UCL), London, UK.* (dic. 2010).

[22] Niccolò Di Marco et al. «Patterns of linguistic simplification on social media platforms over time». In: *Proceedings of the National Academy of Sciences of the United States of America* 121.50 (2024), e2412105121. DOI: 10.1073/pnas.2412105121.

[23] Michael E. W. Varnum et al. «Why are song lyrics becoming simpler? A time series analysis of lyrical complexity in six decades of American popular music». In: *PLoS ONE* 16.1 (2021), e0244576. DOI: 10.1371/journal.pone.0244576.

[24] Janis Libeks e Douglas Turnbull. «You Can Judge an Artist by an Album Cover: Using Images for Music Annotation». In: *Multimedia, IEEE* 18 (mag. 2011), pp. 30–37. DOI: 10.1109/MMUL.2011.1.

[25] Nathan Lee e Robert Baraldi. *Predicting Musical Genre from Album Cover Art*. Rapp. tecn. URL: https://cs230.stanford.edu/projects_

spring_2020/reports/38953572.pdf. Stanford University, CS230 Project, 2020.

[26] Aleksandra Dorochowicz e Bozena Kostek. «Relationship between album cover design and music genres». In: set. 2019, pp. 93–98. DOI: 10.23919/SPA.2019.8936738.

[27] Kaelen Greenfield e Ebenezer Paintsil. «Album Covers Deserve Some Attention». In: (2024). URL: https://cs231n.stanford.edu/2024/papers/album-covers-deserve-some-attention.pdf.

[28] Tara Venkatesan, Qian Wang e Charles Spence. «Does the Typeface on Album Cover Influence Expectations and Perception of Music?» In: *Psychology of Aesthetics, Creativity, and the Arts* 16 (ago. 2020), pp. 487–503. DOI: 10.1037/aca0000330.

[29] Maria Grebenkina et al. «Edge-Orientation Entropy Predicts Preference for Diverse Types of Man-Made Images». In: *Frontiers in Neuroscience* 12 (2018), p. 678. DOI: 10.3389/fnins.2018.00678.

[30] Sergio Oramas et al. «Multi-label Music Genre Classification from audio, text and images using Deep Features». In: *Proceedings of the 18th International Society for Music Information Retrieval Conference (ISMIR 2017)*. 2017. URL: https://arxiv.org/abs/1707.04916.

[31] Thierry Bertin-Mahieux et al. «The Million Song Dataset». In: *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR)*. Miami, FL, USA, 2011, pp. 591–596.

[32] MetaBrainz Foundation. *MusicBrainz*. https://musicbrainz.org/. Accesso: 2024-07-01.

[33] Sergio Oramas et al. «Multimodal Deep Learning for Music Genre Classification». In: *Transactions of the International Society for Music Information Retrieval* 1.1 (2018). DOI: 10.5334/tismir.13.

[34] AcousticBrainz. *Million Song Dataset to MusicBrainz mapping*. https://labs.acousticbrainz.org/million-song-dataset-mapping/. Accesso: 2024-07-01.

[35] Allen Guo. *billboard.py: A Python API for scraping billboard.com charts*. https://github.com/guoguo12/billboard-charts. Accesso: 2025-07-14. 2014.

[36] Musicmap. *The Genealogy and History of Popular Music Genres*. https://musicmap.info. Accesso: Luglio 2025. 2024.

[37] Google. *Gemini: The most capable model from Google*. https://deepmind.google/technologies/gemini/. Modello utilizzato: Gemini 2.5 Pro (Maggio 2025). Accesso: Luglio 2025. 2025.

[38] Google. *Gemini 2.0 Flash: Model Card*. https://storage.googleapis.com/model-cards/documents/gemini-2-flash.pdf. Modello utilizzato: gemini-2.0-flash. Accesso: Maggio 2025. 2025.

[39] Glenn Jocher et al. *YOLOv8*. https://github.com/ultralytics/ultralytics. Accesso: 14 Luglio 2025. 2023.

[40] Wikipedia contributors. *Record label*. https://en.wikipedia.org/wiki/Record_label. Accesso: 14 Luglio 2025. 2025.