UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
CURSO DE ENGENHARIA DA COMPUTAÇÃO

JÉSSICA AMARAL ALMEIDA ALENCAR DE OLIVEIRA

FORMATAÇÃO E SINCRONIZAÇÃO DE CIFRAS: um sistema baseado em inteligência artificial

Recife

2025

JÉSSICA AMARAL ALMEIDA ALENCAR DE OLIVEIRA

FORMATAÇÃO E SINCRONIZAÇÃO DE CIFRAS: um sistema baseado em inteligência artificial

Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia da Computação da Universidade Federal de Pernambuco, Centro de Informática, como requisito para a obtenção do título de Bacharel em Engenharia da Computação.

Orientador: Filipe Calegario

Recife

2025

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Oliveira, Jéssica Amaral Almeida Alencar de.
Formatação e Sincronização de Cifras: Um Sistema Baseado em Inteligência Artificial / Jéssica Amaral Almeida Alencar de Oliveira. - Recife, 2025.
50 p. : il., tab.

Orientador(a): Filipe Calegario
Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Informática, Engenharia da Computação - Bacharelado, 2025.
Inclui referências, apêndices.

1. Alinhamento acorde-letra. 2. Construção de cifras. 3. Music Information Retrieval (MIR). 4. Ferramentas para músicos. 5. Cifragem automatizada. I. Calegario, Filipe. (Orientação). II. Título.

000 CDD (22.ed.)

JÉSSICA AMARAL ALMEIDA ALENCAR DE OLIVEIRA

FORMATAÇÃO E SINCRONIZAÇÃO DE CIFRAS:

um sistema baseado em inteligência artificial

Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia da Computação da Universidade Federal de Pernambuco, Centro de Informática, como requisito para a obtenção do título de Bacharel em Engenharia da Computação.

Aprovado em: 09/04/2025.

BANCA EXAMINADORA

_____

Profº. Dr. Filipe Calegario (Orientador)

Universidade Federal de Pernambuco

_____

Profº. Dr. Giordano Cabral (Examinador Interno)

Universidade Federal de Pernambuco

Dedico este trabalho à minha mãe, por seu amor incondicional e por sempre acreditar em mim, mesmo nos momentos em que duvidei de mim mesma. Ao meu pai (In Memoriam), que sempre sonhou com a minha formatura e, de onde estiver, sei que se orgulha desta conquista. E a todos que, com palavras, gestos e apoio, me deram forças para continuar.

AGRADECIMENTOS

Agradeço ao professor Filipe Calegario, que aceitou embarcar nessa jornada comigo, me guiando e direcionando até a conclusão deste trabalho, e por, junto com o professor Giordano Cabral, mostrar que a universidade também pode ser um lugar de aprendizado leve e divertido.

Aos professores Fernando Castor e Carlos Alexandre (famoso CABM), minha profunda gratidão pela demonstração de empatia e humanidade em um ambiente que, muitas vezes, pode ser sufocante. Suas atitudes não passaram despercebidas e fizeram toda a diferença.

À equipe da Apple Developer Academy e à minha turma de 2020, meu sincero agradecimento. Foram dois anos desafiadores, mas ter vocês ao meu lado tornou tudo mais leve. Foi graças a esse projeto que minha carreira encontrou um caminho promissor.

E, por fim, a todos os colegas que compartilharam essa jornada comigo, obrigada por cada troca, cada apoio e cada risada no meio do caos. Foi por causa de vocês que consegui chegar até aqui sem pirar completamente.

RESUMO

Este trabalho propõe um sistema automatizado para formatação e sincronização de cifras musicais a partir de arquivos de áudio e letras, utilizando técnicas de processamento musical e inteligência artificial. O estudo concentra-se na organização estrutural das cifras no padrão brasileiro, desenvolvendo métodos para alinhar temporalmente acordes extraídos do áudio com as palavras correspondentes na letra da música. O sistema integra ferramentas de Music Information Retrieval (MIR) para extração automática de acordes com técnicas de force alignment para sincronização texto-áudio. Um dos principais desafios foi desenvolver métodos eficientes para posicionar os acordes corretamente sobre as palavras, mantendo a estrutura rítmica e harmônica da música. Os resultados mostram que, embora a extração de acordes atinja 47% de precisão em comparação com cifras de referência, o alinhamento temporal ainda apresenta limitações, com apenas 2,64% dos acordes posicionados de forma ideal. A principal contribuição deste trabalho está na automatização do processo de formatação de cifras, que tradicionalmente é manual e propenso a erros. O sistema oferece um modelo viável para aplicações práticas, como educação musical e ferramentas de aprendizado, enquanto aponta caminhos para melhorias futuras, como o refinamento do alinhamento temporal e a inclusão de variações harmônicas mais complexas.

Palavras-chaves: Alinhamento acorde-letra. Music Information Retrieval (MIR). Formatação de cifras. Ferramentas para músicos. Cifragem automatizada.

ABSTRACT

This work proposes an automated system for formatting and synchronizing musical chord charts from audio files and lyrics, employing music processing techniques and artificial intelligence. The study focuses on the structural organization of chord charts following Brazilian standards, developing methods to temporally align chords extracted from audio with corresponding lyrics.

The system integrates Music Information Retrieval (MIR) tools for automatic chord extraction with force alignment techniques for text-audio synchronization. A key challenge involved developing efficient methods to correctly position chords over lyrics while preserving the musical harmonic and rhythmic structure. Results show that while chord extraction achieves 47% accuracy compared to reference charts, temporal alignment remains limited, with only 2.64% of chords optimally positioned.

The main contribution lies in automating the chord chart formatting process, which is traditionally manual and error-prone. The system provides a viable model for practical applications in music education and learning tools, while suggesting future improvements such as refining temporal alignment and incorporating more complex harmonic variations.

Keywords: Chord-lyrics alignment. Music Information Retrieval (MIR). Chord chart formatting. Tools for musicians. Automated chord notation.

LISTA DE FIGURAS

Figura 1 – Formato dos Acordes Extraídos 22
Figura 2 – Formato das Letras Extraídas 22
Figura 3 – Estrutura do Sistema 25
Figura 4 – Letra extraída do *Letras.mus*. 26
Figura 5 – Fluxo do Lyrics Sync 26
Figura 6 – Exemplo das dificuldades no alinhamento entre acordes e letra. 28
Figura 7 – Resultado do processamento das cifras para validação. 28
Figura 8 – Métricas gerais comparando cifras geradas e de validação. 35
Figura 9 – Distribuição dos tipos de edição identificados nas cifras geradas. 37
Figura 10 – Acordes mais frequentes presentes exclusivamente em cada cifra. 38

LISTA DE CÓDIGOS

Código Fonte 1 – Alinhamento de Acorde Sobre Palavra 32

LISTA DE TABELAS

Tabela 1 – Regras de tradução de nomenclatura 27
Tabela 2 – Inversões de acordes e suas variantes 48
Tabela 3 – Lista de músicas com URLs de áudio, letra e cifra. 51

LISTA DE ABREVIATURAS E SIGLAS

ACE	Automatic Chord Estimation
ISMIR	International Society for Music Information Retrieval
MIR	Music Information Retrieval

SUMÁRIO

1 INTRODUÇÃO 14

2 FUNDAMENTAÇÃO TEÓRICA 17

2.1 Music Information Retrieval (MIR) E EXTRAÇÃO AUTOMÁTICA DE ACORDES 17

2.2 REPRESENTAÇÃO MUSICAL: CIFRAS VS. TABLATURAS 17

2.3 ALINHAMENTO TEMPORAL E INTEGRAÇÃO MULTIMODAL 18

2.4 LACUNAS E ORIGINALIDADE DO PRESENTE TRABALHO 18

3 OBJETIVOS 20

3.1 OBJETIVO GERAL 20

3.2 OBJETIVOS ESPECÍFICOS 20

4 METODOLOGIA 21

4.1 FERRAMENTAS UTILIZADAS 21

4.2 FLUXO DE DESENVOLVIMENTO 22

4.2.1 Extração de Áudio 23

4.2.2 Extração de Acordes 23

4.2.3 Recuperação de Letra 23

4.2.4 Extração de Letra Sincronizada 23

4.2.5 Alinhamento da Letra com os Acordes 23

4.2.6 Recuperação de Cifras para Validação 24

4.2.7 Comparação e Extração de Métricas 24

5 DESENVOLVIMENTO E DESAFIOS 25

5.1 ETAPAS DO DESENVOLVIMENTO 25

5.1.1 Extração do Áudio 25

5.1.2 Recuperação da Letra da Música 26

5.1.3 Extração da Letra Sincronizada 26

5.1.4 Extração dos Acordes 27

5.1.5 Alinhamento de Acordes e Formatação da Cifra 27

5.1.6 Recuperação e Processamento de Cifras para Validação 28

5.1.7 Avaliação dos Resultados 29

5.2 DESAFIOS E SOLUÇÕES 29

5.2.1 Desafios na Extração de Áudio . . . 29
5.2.2 Extração de Letras e Alternativas Consideradas . . . 29
5.2.3 Uso de Force Alignment para Alinhamento da Letra . . . 30
5.2.4 Extração de Acordes e Limitações das Ferramentas . . . 30
5.2.5 Conflitos entre Pacotes e Ajustes na Ordem de Execução . . . 31
5.2.6 Alinhamento dos Acordes com a Letra . . . 31
5.2.7 Ajustes para Avaliação dos Resultados . . . 32
6 AVALIAÇÃO . . . 33
6.1 CONSTRUÇÃO DA BASE DE DADOS . . . 33
6.2 MÉTODOS DE AVALIAÇÃO . . . 34
6.3 ANÁLISE DOS RESULTADOS . . . 35
6.3.1 Métricas Gerais . . . 35
6.3.2 Análise das Sequências de Acordes . . . 37
6.3.3 Acordes Mais Frequentes em Cada Cifra . . . 38
7 DISCUSSÃO . . . 39
7.1 SIMPLIFICAÇÃO DOS ACORDES . . . 39
7.2 CRITÉRIO DE POSICIONAMENTO DOS ACORDES . . . 39
7.3 CRITÉRIO DE CORRESPONDÊNCIA DOS ACORDES . . . 39
7.4 VARIAÇÃO ENTRE VERSÕES DA MESMA MÚSICA . . . 40
7.5 ANÁLISE QUANTITATIVA VS. QUALITATIVA . . . 40
7.6 DEPENDÊNCIA DA CIFRA DE VALIDAÇÃO . . . 40
8 CONCLUSÃO . . . 41
8.1 DIREÇÕES PARA MELHORIAS FUTURAS . . . 41
REFERÊNCIAS . . . 43
APÊNDICE A – DICIONÁRIO DE INVERSÕES DE ACORDES . . . 46
APÊNDICE B – BASE DE DADOS . . . 49

1 INTRODUÇÃO

O violão é amplamente difundido no Brasil, permeando todas as classes sociais. Márcia Taborda afirma que ele continua sendo um instrumento que sempre se difundiu por sua portabilidade, facilidade e preço, não só porque está relacionado aos diversos gêneros da MPB, mas pelas facilidades que implica, apesar de ter uma técnica muito difícil de ser dominada. (TABORDA, 2011)

Historicamente, os métodos de ensino do violão remontam às antigas tablaturas publicadas desde o século XVI, que visavam tornar mais acessível a literatura do instrumento. Com o tempo, essas tablaturas evoluíram para um método conhecido como cifragem (TABORDA, 2011). Almir Chediak define a cifragem como a utilização de símbolos criados para representar acordes de maneira prática, composta por letras, números e sinais, sendo o sistema predominantemente usado na música popular para qualquer instrumento (CHEDIAK, 1986). Graças à cifragem, o aprendizado do violão tornou-se mais prático e simples, resultando na impressão de métodos baseados nesse sistema em maior quantidade em comparação aos que utilizam a notação musical tradicional (TABORDA, 2011).

Com o avanço tecnológico, surgiram plataformas digitais dedicadas à disseminação de cifras, como o Cifra Club (SOL, 2025a). Inicialmente criado para compartilhar cifras de violão, o Cifra Club transformou-se em uma das principais plataformas dedicadas aos amantes da música no Brasil. Essas plataformas frequentemente contam com o apoio da comunidade de músicos para a construção e compartilhamento de cifras, o que, embora promova uma vasta coleção de músicas, pode resultar em variações na qualidade e precisão das cifras disponibilizadas.

Atualmente, existem ferramentas que utilizam inteligência artificial para identificar acordes de músicas de forma automática, como o *Chord AI* (OU, 2025), que detecta e exibe acordes em sincronia com qualquer música, permitindo transposição em tempo real e ajuste de complexidade dos acordes. Outra ferramenta notável é o *Moises* (INC., 2025), um aplicativo que oferece separação de áudio por meio de inteligência artificial, permitindo isolar vocais e instrumentos, além de detectar e exibir acordes em tempo real. O *Moises* também possibilita a transposição de tons e ajuste de velocidade das músicas, facilitando a prática musical para instrumentistas e vocalistas.

No entanto, a geração automática e a sincronização de cifras a partir de letras e fonogramas (áudio das músicas) apresentam desafios significativos. A precisão na detecção de acordes

e o alinhamento correto com a letra são dificuldades técnicas que impactam diretamente a usabilidade e confiabilidade dessas ferramentas. A implementação de sistemas automatizados capazes de gerar cifras precisas e sincronizadas é de grande relevância para músicos e educadores musicais, pois proporciona economia de tempo e maior precisão no ensino e na prática musical.

Além disso, é relevante destacar as diferenças na notação musical entre o Brasil e outros países. Enquanto no Brasil o uso de cifras é predominante, especialmente entre violonistas, em muitos países estrangeiros, como os de língua inglesa, a tablatura é mais comumente utilizada. A tablatura fornece uma representação gráfica das cordas e casas do instrumento, indicando exatamente onde os dedos devem ser posicionados. Essa distinção reflete diferenças culturais e pedagógicas no ensino e prática musical, influenciando a forma como músicos aprendem e compartilham músicas em diferentes regiões.

Portanto, a integração de tecnologias avançadas na geração e sincronização de cifras, aliada à consideração das particularidades culturais na notação musical, representa um avanço significativo no campo da educação musical, atendendo às demandas de uma comunidade musical cada vez mais digital e interconectada.

Nesse contexto, o presente trabalho propõe o desenvolvimento de um sistema inovador que automatiza a formatação e sincronização de cifras a partir de letras e fonogramas, enfrentando desafios como a detecção precisa de acordes e o alinhamento correto com a letra. Nossa abordagem integra ferramentas existentes na área de MIR para facilitar a transcrição de músicas em cifras. Especificamente, utilizamos ferramentas de extração de letras sincronizadas que empregam técnicas avançadas, como separação de fontes, para isolar vocais e instrumentos em gravações musicais. Essas ferramentas permitem obter letras alinhadas temporalmente com o áudio, facilitando a identificação dos momentos exatos em que cada palavra é cantada. Com as letras sincronizadas, aplicamos algoritmos de detecção de acordes para identificar as harmonias correspondentes em cada segmento da música. Posteriormente, desenvolvemos procedimentos para alinhar os acordes detectados com as respectivas palavras na letra, resultando em cifras precisas e sincronizadas. Além disso, implementamos métodos de validação e correção das cifras geradas, garantindo maior precisão e confiabilidade nos resultados. Essa integração de ferramentas e técnicas proporciona uma automatização eficaz na obtenção de cifras sincronizadas, facilitando o aprendizado e a prática musical para instrumentistas e vocalistas.

A estrutura deste trabalho está organizada da seguinte forma: no Capítulo 2, apresentamos a fundamentação teórica, abordando conceitos relacionados à recuperação de informação

musical MIR e extração automática de acordes, bem como a representação musical por cifras e tablaturas. O Capítulo 3 descreve os objetivos do trabalho, detalhando o objetivo geral e os objetivos específicos. No Capítulo 4, detalhamos a metodologia empregada no desenvolvimento do sistema proposto, incluindo as ferramentas e técnicas utilizadas, além do fluxo de desenvolvimento. O Capítulo 5 aborda o desenvolvimento e os desafios enfrentados durante a implementação do sistema. No Capítulo 6, apresentamos a avaliação do sistema, incluindo a construção da base de dados e a análise dos resultados obtidos. O Capítulo 7 traz a discussão dos resultados, abordando aspectos como simplificação dos acordes e critérios de posicionamento. Por fim, no Capítulo 8, apresentamos as conclusões e sugestões para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 MIR E EXTRAÇÃO AUTOMÁTICA DE ACORDES

A área de MIR dedica-se ao desenvolvimento de técnicas computacionais para análise e extração de informações musicais a partir de dados multimodais (SCHEDL; GÓMEZ; URBANO, 2014). Dentre suas subáreas, o Automatic Chord Estimation (ACE) destaca-se por identificar acordes em gravações de áudio, utilizando métodos como:

- Análise espectral (features chroma e MFCCs) para capturar informações harmônicas (MCVICAR et al., 2014);
- Redes neurais (ex.: CNNs, Transformers) para classificação de acordes em tempo real (ROWE; TZANETAKIS, 2021).

Entretanto, como observado na revisão da base de dados da International Society for Music Information Retrieval (ISMIR) (2019–2024), a maioria dos trabalhos em ACE — como em (MICCHI et al., 2021) — limita-se à detecção de acordes isolados, sem integrá-los a letras ou formatá-los em cifras.

2.2 REPRESENTAÇÃO MUSICAL: CIFRAS VS. TABLATURAS

Enquanto a notação musical internacional privilegia tablaturas, como em Edwards et al. (2024), a música popular brasileira adota cifras como padrão, devido à sua simplicidade e disseminação em plataformas como Cifra Club. Cifras representam acordes por:

- Símbolos alfabéticos (ex.: C, Dm, G7);
- Sufixos que indicam qualidade (ex.: "m"= menor, "7"= sétima) e inversões (ex.: C/E) (DOBBERT, 2020).

Apesar dessa relevância cultural, nenhum estudo na ISMIR aborda a automatização na construção e sincronização de cifras, evidenciando uma lacuna na literatura.

2.3 ALINHAMENTO TEMPORAL E INTEGRAÇÃO MULTIMODAL

O alinhamento preciso entre eventos harmônicos e textuais permanece um desafio subexplorado em MIR, especialmente no contexto de cifras formatadas. Embora trabalhos recentes como Poltronieri, Presutti e Rocamora (2024) e Gao, Gupta e Li (2022) avancem em sincronização áudio-texto, suas abordagens negligenciam:

- Formatação prática de cifras (ex.: blocos sobre letras, como no Cifra Club);
- Adaptação a gêneros brasileiros, onde variações rítmicas (ex.: samba) e harmônicas (ex.: acordes extendidos em MPB) exigem tratamento especializado.

2.4 LACUNAS E ORIGINALIDADE DO PRESENTE TRABALHO

Nos últimos cinco anos, a pesquisa em extração de acordes a partir de áudio tem avançado significativamente. Um exemplo notável é o sistema *DECIBEL*, apresentado por Odekerken, Koops e Volk (2021), que integra representações simbólicas heterogêneas, como arquivos *MIDI* e tablaturas, para aprimorar a estimação automática de acordes (ACE). O *DECIBEL* alinha arquivos *MIDI* e tablaturas com gravações de áudio, utilizando técnicas específicas de cada representação para estimar sequências de acordes, resultando em uma representação harmônica rica que melhora os métodos tradicionais de ACE baseados apenas em áudio.

Paralelamente, a formatação automática de tablaturas também recebeu atenção. Em 2023, Dalmazzo, Déguernel e Sturm (2023) introduziram o *Chordinator*, um modelo baseado em transformadores treinado com um extenso conjunto de sequências de acordes de diversos estilos musicais. Eles exploraram diferentes estratégias de modelagem, incluindo a "tokenização"de acordes como elementos únicos e a expressão de acordes como tuplas descrevendo a fundamental, natureza e extensões. A terceira estratégia incorporou informações de notas *MIDI* às *embeddings* posicionais, resultando em progressões de acordes mais consistentes e musicalmente plausíveis.

Além disso, a transcrição de letras de músicas tem sido um foco de pesquisa, especialmente com o advento de ferramentas avançadas de transcrição de texto. No entanto, apesar desses avanços individuais em extração de acordes, formatação de tablaturas e transcrição de letras, não foram encontrados trabalhos que integrem essas ferramentas para automatizar o processo completo de formatação de cifras. A conferência ISMIR, sendo uma das mais importantes na

área de Recuperação de Informação Musical, reflete as tendências e lacunas atuais no campo. A ausência de pesquisas que combinem essas áreas indica uma oportunidade para desenvolver um *pipeline* que una a extração de acordes a partir de áudio, alinhamento temporal com letras e formatação automática em padrões de cifras, atendendo a uma demanda real de músicos brasileiros.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Desenvolver um sistema capaz de automatizar a construção de cifras musicais a partir da letra e do áudio de uma música, garantindo precisão e padronização no processo.

3.2 OBJETIVOS ESPECÍFICOS

- Identificar e utilizar ferramentas para analisar músicas, extrair letras e acordes a partir do áudio, automatizando o processo de formatação de cifras.
- Garantir que as cifras geradas sejam compatíveis com os formatos mais utilizados por músicos brasileiros.
- Sincronizar as cifras com a letra da música para facilitar sua execução.
- Testar e validar o sistema com diferentes cifras presentes em sites amplamente difundidos entre músicos.

4 METODOLOGIA

4.1 FERRAMENTAS UTILIZADAS

O projeto foi desenvolvido utilizando diversas ferramentas e tecnologias para garantir a eficiência e a precisão dos resultados. As principais ferramentas empregadas foram:

- *Google Colab* (GOOGLE, 2025): Utilizado para o desenvolvimento e execução do código, aproveitando o suporte a *GPUs* para acelerar os processamentos mais exigentes.
- *GitHub* (GITHUB, 2025): Utilizado para versionamento do código e para garantir a segurança e integridade dos dados ao longo do desenvolvimento.
- *Python* (FOUNDATION, 2025c): Linguagem de programação principal do projeto, devido à sua vasta biblioteca para manipulação de áudio, *web scraping* e análise de dados.

Para a extração de áudios das músicas, foi utilizado o *YouTube* (YOUTUBE, 2025) como fonte, em conjunto com a ferramenta de linha de comando *yt-dlp*, devido à sua confiabilidade e à experiência prévia no seu uso. O *yt-dlp* recebe o link do vídeo como entrada e retorna o áudio no formato desejado.

Para a extração dos acordes, utilizou-se a *Large Vocabulary Chord Recognition* (JIANG et al., 2019), apresentada na conferência ISMIR, em 2019. Esta ferramenta oferece uma ampla diversidade de acordes, incluindo:

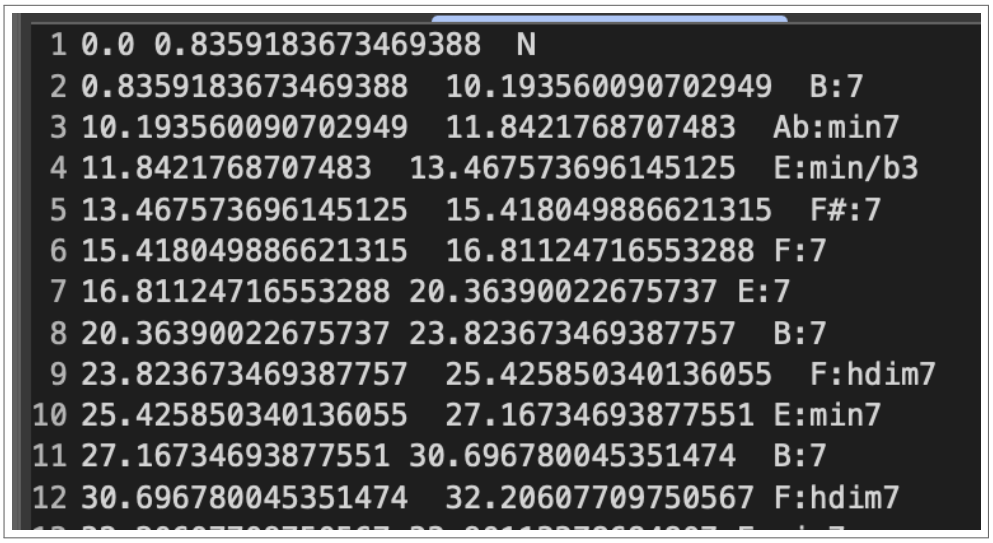
- Tríades básicas e invertidas;
- Acordes com sétima;
- Acordes estendidos e suspensos;
- Acordes com baixo diferente.

A ferramenta recebe o áudio como entrada e retorna uma lista de acordes com *timestamps*, como mostrado na Figura 1.

Para a obtenção de letras de entrada e cifras para validação, foi utilizado *web scraping*, com o auxílio das bibliotecas *BeautifulSoup* (RICHARDSON, 2015) e *requests* (REITZ, 2024).

A extração da letra sincronizada com o áudio foi realizada com o *LyricsSync* (ZHOU, 2023), que recebe o áudio como entrada e retorna as letras com *timestamps* em um arquivo *CSV* (Figura 2).

Na etapa de validação, foi utilizada a biblioteca *Pychord* (MIHIRA, 2023) para a normalização dos acordes. Outras bibliotecas foram utilizadas para comparação das cifras formatadas com as cifras de referência e análise dos resultados, incluindo *difflib* (FOUNDATION, 2025b), *collections* (FOUNDATION, 2025a) e *numpy* (OLIPHANT, 2024).



```
 1 0.0 0.8359183673469388  N
 2 0.8359183673469388  10.193560090702949  B:7
 3 10.193560090702949  11.8421768707483  Ab:min7
 4 11.8421768707483  13.467573696145125  E:min/b3
 5 13.467573696145125  15.418049886621315  F#:7
 6 15.418049886621315  16.81124716553288 F:7
 7 16.81124716553288 20.36390022675737 E:7
 8 20.36390022675737 23.823673469387757  B:7
 9 23.823673469387757  25.425850340136055  F:hdim7
10 25.425850340136055  27.16734693877551 E:min7
11 27.16734693877551 30.696780045351474  B:7
12 30.696780045351474  32.20607709750567 F:hdim7
```

Figura 1 – Formato dos Acordes Extraídos

label	start	end
É	4.18	7.5200000000000005
pau,	7.5600000000000005	7.94
é	8.34	8.38
pedra,	8.42	8.82
é	8.86	8.88
o	8.96	9.24
fim	9.26	9.52
do	9.540000000000001	9.68
caminho	9.700000000000001	10.3
É	10.42	10.56

Figura 2 – Formato das Letras Extraídas

4.2 FLUXO DE DESENVOLVIMENTO

O desenvolvimento do sistema seguiu uma abordagem modular, composta por etapas bem definidas para extração, processamento e validação dos dados musicais. O fluxo do processo pode ser dividido nos seguintes passos:

4.2.1 Extração de Áudio

Inicialmente, o fonograma da música é extraído e convertido para um formato adequado ao processamento digital. Essa etapa garante que o áudio esteja em uma qualidade compatível com as ferramentas de análise utilizadas nas etapas subsequentes.

4.2.2 Extração de Acordes

Após a extração do áudio, um algoritmo de reconhecimento harmônico é aplicado para identificar os acordes musicais presentes na gravação. Os acordes extraídos são então convertidos para formatos mais comuns utilizados em cifras brasileiras, garantindo a padronização da saída do sistema.

4.2.3 Recuperação de Letra

Paralelamente à extração dos acordes, a letra da música é obtida de fontes online ou de bases de dados de letras. Esse processo é fundamental para viabilizar o alinhamento posterior entre a letra e os acordes.

4.2.4 Extração de Letra Sincronizada

Com a letra recuperada e o áudio extraído, é aplicada uma ferramenta de *force alignment* para sincronizar as palavras da música com o fonograma. Esse alinhamento temporal permite a correspondência exata entre os momentos de execução do áudio e a respectiva letra.

4.2.5 Alinhamento da Letra com os Acordes

Após obter a letra sincronizada e os acordes extraídos, realiza-se a sobreposição dessas informações, garantindo a correta formatação da cifra final. O alinhamento ocorre com base nos tempos identificados na extração de acordes e na segmentação da letra.

4.2.6 Recuperação de Cifras para Validação

Para garantir a qualidade e a confiabilidade do sistema, cifras já disponíveis em fontes confiáveis são recuperadas e utilizadas como referência para avaliação dos resultados.

4.2.7 Comparação e Extração de Métricas

Por fim, as cifras geradas pelo sistema são comparadas com as cifras de referência por meio de métricas quantitativas e qualitativas. Essa etapa busca validar a precisão do alinhamento entre acordes e letra, além de aferir a fidelidade da saída em relação às cifras tradicionais.

5 DESENVOLVIMENTO E DESAFIOS

5.1 ETAPAS DO DESENVOLVIMENTO

O desenvolvimento do sistema foi dividido em várias etapas, desde a extração do áudio até a formatação final da cifra. A seguir, cada uma dessas fases é detalhada.

A Figura a seguir ilustra a arquitetura do sistema desenvolvido.[1]

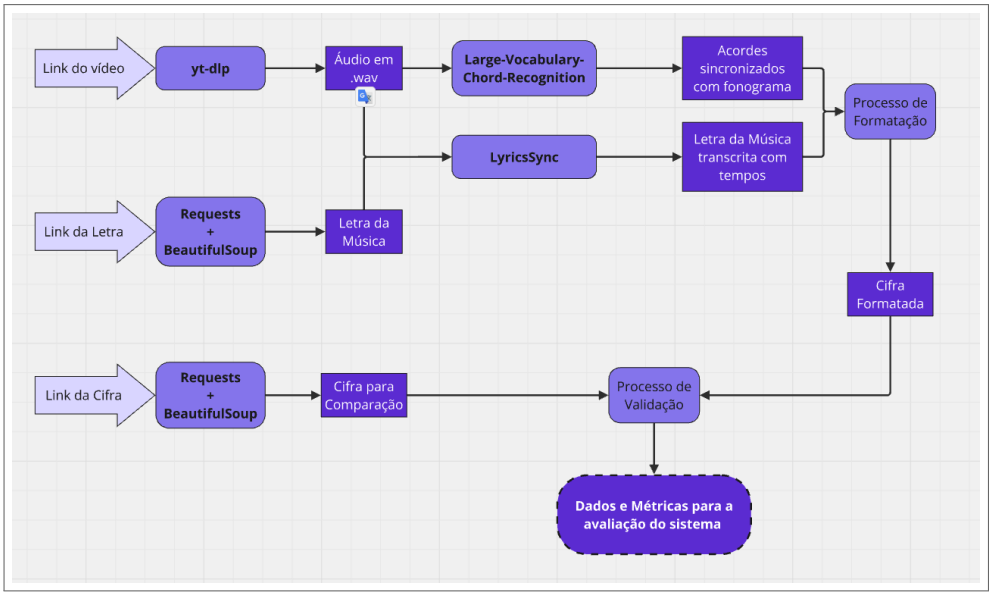


Figura 3 – Estrutura do Sistema

5.1.1 Extração do Áudio

A primeira decisão tomada foi utilizar o *YouTube* (YOUTUBE, 2025) como fonte de extração do áudio, devido à grande diversidade de músicas disponíveis. Após uma análise das ferramentas disponíveis, foi escolhida a ferramenta *yt-dlp* (CONTRIBUTORS, 2025), uma ferramenta de linha de comando amplamente utilizada para esse tipo de tarefa e com a qual já havia familiaridade. O processo inicial foi testado com apenas uma música, e o *dataset* oficial só foi montado ao final do desenvolvimento. O formato de saída do áudio foi definido como .wav, devido à sua ampla compatibilidade com a maioria das ferramentas utilizadas nos testes.

[1] O código-fonte está disponível em: <https://github.com/jessica-aaao/AutoCifra>

5.1.2 Recuperação da Letra da Música

Para a obtenção das letras, inicialmente considerou-se o uso da *Genius API* (Genius API, 2025). No entanto, como o foco do projeto era o público brasileiro, optou-se por uma fonte que proporcionasse maior disponibilidade de letras em português. Assim, foi selecionado o site *Letras.mus* (SOL, 2025b), sendo empregada a técnica de *web scraping* para a extração automatizada das letras. A Figura 4 apresenta um exemplo do resultado obtido a partir desse processo.

```
21  Little things I should have said and done
22  I just never took the time
23  You were always on my mind
24  You were always on my mind
25  You were always on my mind
```

Figura 4 – Letra extraída do *Letras.mus*.

5.1.3 Extração da Letra Sincronizada

Essa etapa foi uma das mais demoradas. Inicialmente, buscou-se ferramentas específicas para extração de letras a partir do fonograma, mas a maioria das soluções encontradas estava focada em *transcrição de áudio baseada em inteligência artificial*, o que não atendia plenamente ao objetivo do projeto.

Foi então que se encontrou o *LyricsSync* (ZHOU, 2023), uma ferramenta de *force alignment*. Essa ferramenta recebe o áudio e a letra já disponíveis e utiliza o *Demucs* (ROUSSET, 2021) para separar os vocais do restante da faixa. A partir da extração de grafemas e fonemas tanto do áudio quanto da letra, a ferramenta retorna a letra sincronizada e alinhada. A Figura 5 ilustra o funcionamento geral do Lyrics Sync.

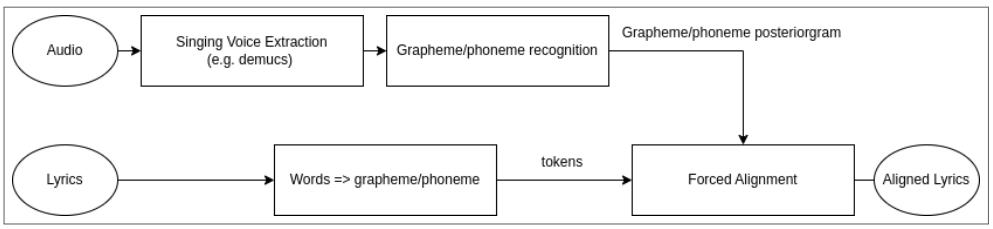


Figura 5 – Fluxo do Lyrics Sync

5.1.4 Extração dos Acordes

Para a extração dos acordes diretamente do fonograma, foi utilizado o *Chord Recognition* (JIANG et al., 2019). O arquivo *.lab* retornado pela ferramenta foi analisado, e algumas transformações foram realizadas para adequar os nomes dos acordes ao formato mais comum no Brasil. Essas transformações podem ser observadas na Tabela 1 e no Anexo A.

Original	**Substituição**
`:`	(removido)
`min`	m
`maj`	(removido)
`hdim7`, `hdim`	m7(b5)
`sus4(b7)`	7sus4

Tabela 1 – Regras de tradução de nomenclatura

5.1.5 Alinhamento de Acordes e Formatação da Cifra

Um dos maiores desafios do desenvolvimento foi o *alinhamento dos acordes* com as palavras da letra, garantindo a correta *formatação da cifra*. Para isso, a maior parte do processo foi realizada de forma manual. Inicialmente, foi necessário comparar os tempos dos acordes com os tempos de cada palavra, a fim de posicionar e alinhar corretamente os acordes em relação à letra.

Esse processo envolveu identificar a posição do acorde em relação às palavras da música, ou seja, se o acorde estava antes, entre ou sobre as palavras. Além disso, foi realizada uma análise proporcional dos tempos, para definir com maior precisão a posição exata do acorde dentro da palavra, considerando o ritmo e a duração de cada sílaba ou palavra da letra. Esse alinhamento foi fundamental para garantir que a cifra final estivesse correta, tanto em termos de harmonia quanto de visualização.

A Figura 6 ilustra visualmente essa dificuldade, mostrando como pequenas variações no posicionamento dos acordes podem impactar a leitura e a interpretação da cifra. As diferenças entre as duas versões destacam os desafios enfrentados durante o processo de alinhamento.

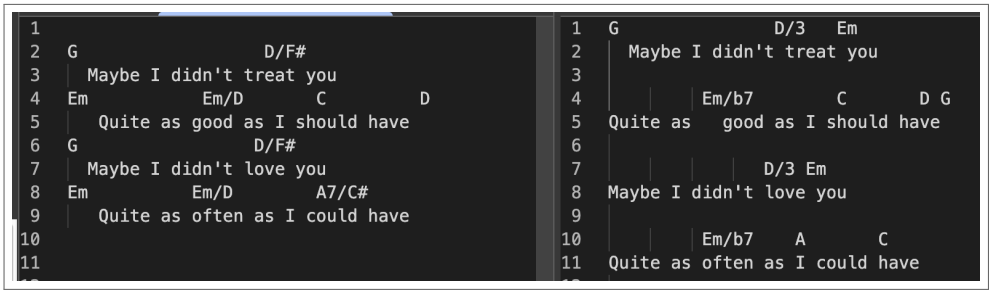


Figura 6 – Exemplo das dificuldades no alinhamento entre acordes e letra.

5.1.6 Recuperação e Processamento de Cifras para Validação

A penúltima etapa do processo foi a *recuperação de cifras para a avaliação do desempenho do sistema*. Para essa tarefa, adotou-se a mesma estratégia usada na obtenção das letras, aplicando-se *web scraping* para extrair cifras do *Cifra Club* (SOL, 2025a), um dos sites mais populares entre músicos, especialmente violonistas.

Após a recuperação das cifras, foi necessário realizar um processamento adicional para remover elementos desnecessários, como tablaturas, textos explicativos e cifras isoladas (como introdução e *bridge*). O objetivo foi tornar as cifras extraídas mais próximas da saída do sistema, facilitando a comparação. A Figura 7 ilustra o resultado final após a remoção dos elementos indesejados, evidenciando cifras mais limpas e compatíveis com a saída esperada do sistema.

(a) Cifra antes da remoção dos elementos extras.

(b) Cifra após a remoção dos elementos extras.

Figura 7 – Resultado do processamento das cifras para validação.

5.1.7 Avaliação dos Resultados

O último passo consistiu na avaliação dos resultados obtidos, o que será abordado em detalhes na próxima seção.

5.2 DESAFIOS E SOLUÇÕES

Durante o desenvolvimento deste trabalho, diversos desafios foram enfrentados e solucionados por meio de diferentes estratégias. Esta seção detalha os principais obstáculos encontrados e as respectivas soluções adotadas.

5.2.1 Desafios na Extração de Áudio

O primeiro desafio encontrado foi durante o uso da ferramenta *yt-dlp* (CONTRIBUTORS, 2025). Devido às tentativas do *YouTube* de impedir o uso indevido de seu conteúdo, em algumas execuções a ferramenta solicitava login para acessar determinados vídeos. Para contornar esse problema, foi necessário utilizar os cookies do navegador, adicionando a flag `–cookies`. Apesar dessa solução, em alguns casos foi necessário excluir o ambiente do *colab* e reiniciar o processo para garantir a correta extração do áudio.

5.2.2 Extração de Letras e Alternativas Consideradas

Na fase de extração das letras, diversos desafios surgiram, exigindo a exploração de alternativas. Inicialmente, considerou-se o uso da ferramenta *Demucs* (ROUSSET, 2021) para extrair os vocais e a *Whisper* (OPENAI, 2022) para transcrever o áudio. A ideia era utilizar a letra extraída da web para corrigir possíveis erros de transcrição. No entanto, a correção manual revelou-se extremamente trabalhosa, devido a:

- Diferenças na divisão das linhas entre a transcrição e a letra original.
- Falhas no reconhecimento de conjunções e pontuações.
- Necessidade de modificar o *JSON* com os tempos das palavras para adequá-lo ao formato desejado.

Ferramentas como *difflib* foram testadas para auxiliar na correção automática, mas os resultados não foram satisfatórios.

5.2.3 Uso de Force Alignment para Alinhamento da Letra

Após extensa pesquisa, identificou-se que uma abordagem mais eficiente seria o uso de uma ferramenta de *force alignment*, que realiza automaticamente o alinhamento da letra com o áudio. Inicialmente, tentou-se utilizar o *Aeneas* (READBEYOND,), porém dificuldades na instalação impediram seu uso. Então, optou-se pelo *Lyrics Sync* (ZHOU, 2023), que se mostrou mais adequado ao projeto.

O uso da ferramenta apresentou seus próprios desafios:

- Dependência de um ambiente *Conda*, que não é suportado nativamente pelo *Google Colab*.
- A instalação do *Conda Colab* permitia rodar apenas a base do ambiente, sem possibilidade de criar novos ambientes virtuais.
- A execução precisava ser feita dentro do repositório clonado do projeto.
- O nome dos arquivos de áudio precisava ser ajustado para não conter caracteres especiais, pois isso causava falhas na ferramenta.
- O *Lyrics Sync* não conseguia criar automaticamente as pastas para os *outputs*, sendo necessário criá-las manualmente antes da execução.

Mesmo com esses desafios, o uso da ferramenta permitiu eliminar diversas etapas manuais no desenvolvimento. No entanto, a etapa continuou sendo a mais demorada do processo, levando aproximadamente 3 minutos por música (utilizando *GPU*), com um tempo total de cerca de uma hora para todo o conjunto de dados. A maior parte desse tempo era consumida na extração e alinhamento das letras.

5.2.4 Extração de Acordes e Limitações das Ferramentas

Para a extração dos acordes, foram testadas várias ferramentas. O *Autochord* parecia promissor, mas apresentou diversos pacotes depreciados, impedindo seu uso. O *Essentia* foi

testado, mas não retornava os tempos de cada acorde, o que inviabilizava o alinhamento posterior. O *Madmom* apresentava os mesmos problemas de pacotes desatualizados que o *Autochord*.

Dessa forma, a escolha final foi o *Chord Recognizer*. Embora funcional, a ferramenta utiliza uma nomenclatura de acordes mais comum internacionalmente (`maj`, `min`, `/3`, etc.), o que exigiu a criação de um "dicionário"para traduzir os acordes para a notação mais comum no Brasil. No entanto, essa tradução ainda não abrange todos os tipos de acordes reconhecidos pela ferramenta.

5.2.5 Conflitos entre Pacotes e Ajustes na Ordem de Execução

Outro problema encontrado foi o conflito entre os pacotes utilizados pelo *Chord Recognizer* e pelo *Lyrics Sync*. Quando a extração dos acordes era realizada após a execução do *Lyrics Sync*, a ferramenta apresentava erros, impedindo a extração dos acordes. O problema parecia estar relacionado à incompatibilidade entre bibliotecas carregadas na sessão.

A solução foi alterar a ordem de execução, garantindo que a extração dos acordes ocorresse antes da sincronização da letra.

5.2.6 Alinhamento dos Acordes com a Letra

O alinhamento dos acordes com a letra foi um dos desafios mais complexos. Foi necessário desenvolver um método para associar os tempos dos acordes aos tempos das palavras, garantindo um posicionamento adequado dos acordes na cifra final.

Após diversas tentativas, optou-se por comparar os tempos de início dos acordes com os intervalos das palavras e dos espaços entre elas. Um dos maiores desafios foi lidar com acordes que *iniciavam antes da palavra*, mas se estendiam sobre ela.

Para lidar com esse problema, foi utilizada uma abordagem proporcional:

- A posição do acorde dentro da palavra foi calculada a partir da razão entre a diferença dos tempos de início do acorde e da palavra e a duração total da palavra.
- Essa razão foi então multiplicada pelo tamanho da palavra (em caracteres) para determinar a posição exata do acorde dentro da string da palavra.

- Por fim, a posição foi arredondada e ajustada para garantir que permanecesse dentro dos limites da palavra, evitando índices inválidos.

O código abaixo exemplifica a implementação dessa abordagem:

Código Fonte 1 – Alinhamento de Acorde Sobre Palavra

```
    word_duration = word_end - word_start
    ratio = (chord_start - word_start) / word_duration

    word_len = len(word)
    word_index = int(round(ratio * (word_len - 1)))
    word_index = max(0, min(word_index, word_len - 1))

    chord_position = line_start + word_index
```

5.2.7 Ajustes para Avaliação dos Resultados

Na etapa de avaliação dos resultados, outro problema surgiu: a diferença na formatação entre as cifras geradas pelo sistema e as cifras coletadas para validação.

O principal obstáculo foi a diferença na divisão das linhas, que dificultava a comparação direta entre as cifras. Para contornar esse problema, foram adotadas as seguintes estratégias:

- Identificação e separação das linhas de acordes, permitindo a comparação como uma única linha contínua.
- Remoção de elementos não essenciais das cifras extraídas, como textos de marcação de partes da música, tablaturas e acordes soltos posicionados entre estrofes.

Esses ajustes foram fundamentais para garantir que as métricas extraídas refletissem com precisão o desempenho do sistema em relação às cifras de referência.

6 AVALIAÇÃO

6.1 CONSTRUÇÃO DA BASE DE DADOS

A coleta de dados foi realizada a partir de três fontes principais:

- *YouTube* (YOUTUBE, 2025) – Utilizado para a extração do áudio das músicas selecionadas.
- Letras (SOL, 2025b) – Fonte utilizada para a obtenção das letras das músicas.
- *Cifra Club* (SOL, 2025a) – Base de dados utilizada para a recuperação de cifras, que serviram como referência para validação do sistema.

Para compor o conjunto de dados (*dataset*), foram selecionadas músicas de diferentes estilos, idiomas e níveis de complexidade musical. As músicas escolhidas incluem tanto repertório nacional quanto internacional (inglês), garantindo diversidade de gêneros e variações harmônicas. Além disso, as cifras utilizadas não contêm o uso de capotraste, visando a padronização da análise.

A organização dos dados seguiu uma ordem sequencial, considerando os seguintes critérios:

- Idioma: As músicas foram dispostas conforme o idioma predominante, incluindo português, inglês e bilíngue (português + inglês).
- Complexidade harmônica: As cifras foram ordenadas progressivamente, das mais simples às mais complexas, permitindo uma análise gradual do desempenho do sistema.

Os dados foram organizados em um arquivo no formato *.csv* e armazenados no *Google Drive*, permitindo sua recuperação e manipulação diretamente no *Google Colab* para as etapas de processamento e análise.

O *dataset* final incluiu um total de 20 músicas [1], distribuídas da seguinte forma:

- Músicas nacionais: 13
- Músicas em inglês: 6
- Música bilíngue (português + inglês): 1

[1] A base de dados completa que foi utilizada pode ser vista no Apêndice B.

Essa abordagem garantiu a estruturação adequada dos dados, facilitando a análise automatizada e a validação do sistema.

6.2 MÉTODOS DE AVALIAÇÃO

A etapa de avaliação foi fundamental para medir a eficácia do sistema na extração e formatação de cifras. Para isso, foram comparadas a cifra gerada pelo sistema e a cifra de validação extraída da web. Inicialmente, os acordes foram extraídos de ambas as cifras junto com suas respectivas posições no arquivo, independentemente da disposição por linha. A partir desses dados, foram realizadas diversas análises com o objetivo de mensurar a precisão e a fidelidade do reconhecimento dos acordes e do seu alinhamento com a letra.

As principais análises conduzidas foram:

- Diferença de Quantidade Gerada (%): Comparação entre a quantidade total de acordes reconhecidos e a quantidade presente na cifra de validação. Essa métrica permitiu identificar a média de acordes ausentes ou adicionados indevidamente, fornecendo um indicativo da exatidão da extração.

- Acordes em Comum (%): Mede a proporção de acordes que aparecem em ambas as cifras, independentemente da quantidade de repetições. Isso ajuda a avaliar a eficácia da ferramenta de reconhecimento de acordes na identificação correta das harmonias.

- Só na Cifra Gerada (%) e Só na Cifra de Validação (%): Representam, respectivamente, a quantidade de acordes identificados apenas na cifra gerada e aqueles presentes apenas na cifra de validação. Essas métricas auxiliam na identificação de padrões de erros, como acordes frequentemente omitidos ou inseridos erroneamente pela ferramenta de extração.

- Acordes com Contagem Igual (%): Compara a quantidade de ocorrências de cada acorde entre as cifras. Esse dado é essencial para entender não apenas se os acordes foram corretamente reconhecidos, mas também se sua frequência ao longo da música foi preservada.

- Posições Iguais dos Acordes (%): Mede a precisão do posicionamento dos acordes em relação à letra da música. Essa análise foi fundamental para verificar se a formatação da cifra manteve a coerência rítmica e harmônica esperada.

- Sequências Iguais (%): Avalia se as progressões harmônicas da cifra gerada e da cifra de validação são idênticas. Essa métrica é relevante porque, mesmo que os acordes individuais sejam reconhecidos corretamente, mudanças na ordem ou na repetição podem comprometer a autenticidade da cifra.

Para a análise do posicionamento dos acordes e da similaridade entre sequências, foram utilizadas as formas simplificadas dos acordes, considerando apenas suas raízes. Essa abordagem foi adotada para evitar que variações na nomenclatura (como inversões e extensões) interferissem na avaliação e para identificar casos em que a ferramenta possa ter simplificado acordes mais complexos.

6.3 ANÁLISE DOS RESULTADOS

6.3.1 Métricas Gerais

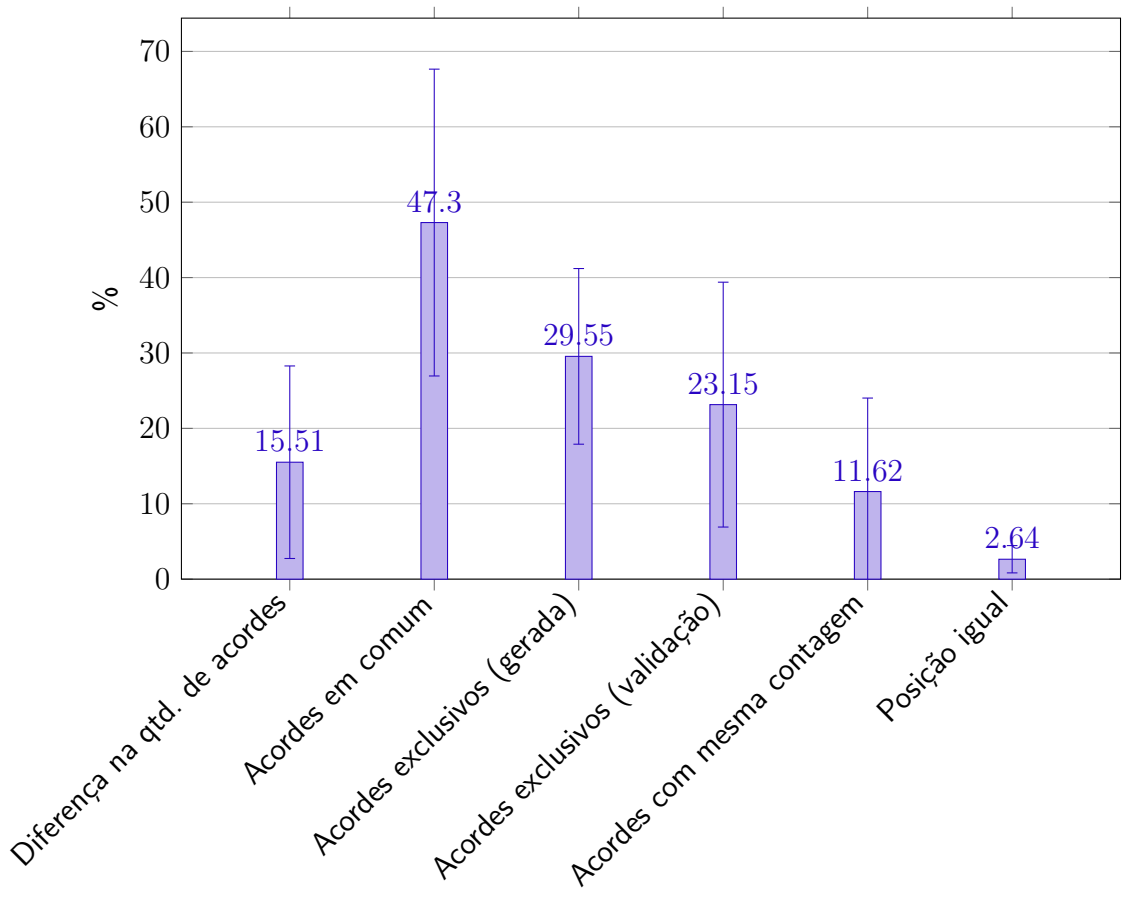


Figura 8 – Métricas gerais comparando cifras geradas e de validação.

A partir do gráfico apresentado acima (Figura 8), observa-se um nível considerável de inconsistência entre as cifras geradas e as cifras de validação. Apesar de uma diferença média de 15,5% na quantidade total de acordes, o alto desvio padrão indica que a ferramenta se comporta de maneira bastante variável entre diferentes músicas. Isso pode significar que, dependendo da estrutura harmônica ou da densidade musical da canção, a ferramenta pode acabar superestimando ou subestimando os acordes identificados.

A média de acordes em comum foi de aproximadamente 47%, sugerindo que, mesmo quando há sobreposição parcial, menos da metade dos acordes coincide exatamente entre as cifras. Isso indica que os erros não estão restritos apenas ao posicionamento, mas também à própria identificação dos acordes.

Outro ponto relevante é o número de acordes exclusivos, tanto nas cifras geradas (29,55%) quanto nas cifras de validação (23,15%). Esses dados mostram que a ferramenta frequentemente acrescenta ou omite acordes, reforçando a necessidade de ajustes no processo de extração.

A métrica de acordes com contagem igual (11,62%) também se destacou por seu baixo valor, indicando dificuldade em captar a frequência correta com que os acordes ocorrem na música. Isso pode estar relacionado à forma como o modelo lida com mudanças sutis na harmonia ou com repetições curtas.

Por fim, a precisão de posicionamento dos acordes em relação à letra foi muito baixa (2,64%). Esse resultado reforça que o alinhamento entre cifra e letra é atualmente o maior desafio, especialmente em casos onde o acorde se estende sobre mais de uma sílaba ou aparece antes do início de uma frase. Esse ponto representa uma oportunidade clara de melhoria para soluções futuras, seja por meio de refinamento nas heurísticas de posicionamento ou pela adoção de novas abordagens baseadas em prosódia e ritmo.

6.3.2 Análise das Sequências de Acordes

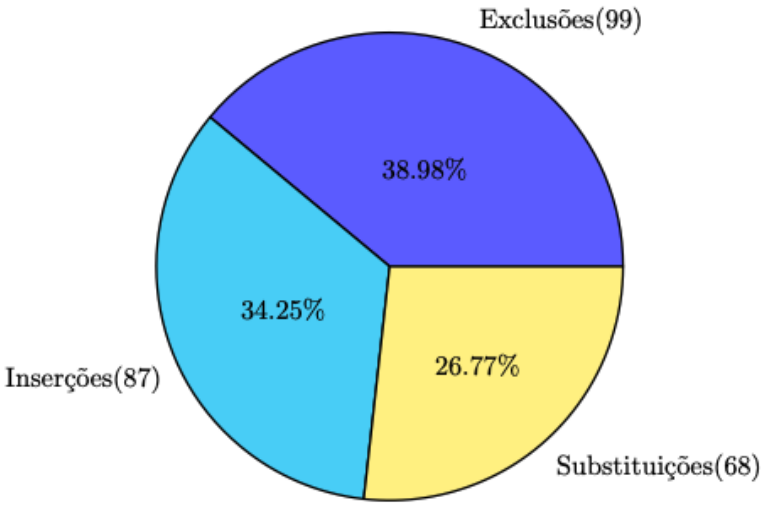


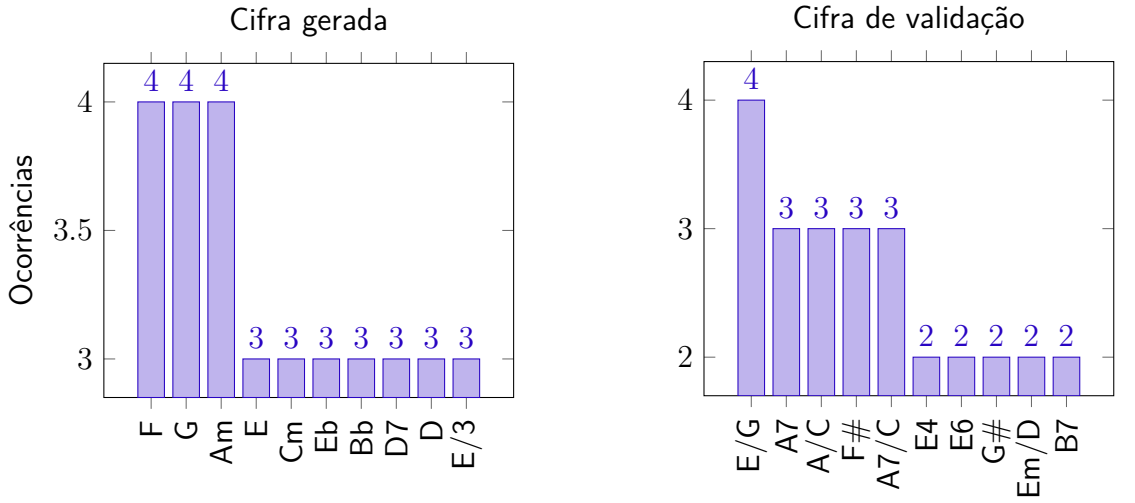
Figura 9 – Distribuição dos tipos de edição identificados nas cifras geradas.

O percentual médio de sequências idênticas entre as cifras foi de 48,4%, com variação pequena entre as músicas. Isso demonstra que, mesmo com erros nos acordes individuais ou no posicionamento, a ferramenta consegue manter boa parte da progressão harmônica principal. Tal resultado indica um comportamento parcialmente consistente quanto à linha harmônica estrutural das músicas.

A análise dos tipos de edição revela que os três tipos de erro (exclusão, inserção e substituição) foram recorrentes: 99 exclusões, 87 inserções e 68 substituições. Embora a exclusão de acordes tenha sido a mais frequente, os demais valores também são expressivos, o que sugere que a ferramenta ainda apresenta falhas relevantes em todas essas dimensões. Isso pode ser reflexo de dificuldades na detecção de acordes sutis, transientes ou com timbres menos destacados.

Essa limitação pode ser explorada em soluções futuras, com a adoção de regras harmônicas mais específicas, detecção de padrões comuns e o uso combinado de diferentes modelos para aumentar a precisão e reduzir inconsistências.

6.3.3 Acordes Mais Frequentes em Cada Cifra



(a) Acordes exclusivos da cifra gerada

(b) Acordes exclusivos da cifra de validação

Figura 10 – Acordes mais frequentes presentes exclusivamente em cada cifra.

A análise dos acordes mais frequentes teve como objetivo investigar possíveis padrões de substituição ou simplificação na identificação dos acordes. No entanto, não foi possível extrair conclusões sólidas a partir desses dados. A presença de acordes como `F`, `G` e `Am` exclusivamente nas cifras geradas pode sugerir uma tendência à simplificação ou superestimação de acordes comuns em tonalidades maiores. Já os acordes exclusivos da cifra de validação, como inversões (`E/G`, `A7/C#`) e extensões (`E4`, `E6`), indicam que a ferramenta tem dificuldade para reconhecer variações harmônicas mais complexas.

Apesar disso, a quantidade relativamente baixa de ocorrências para cada acorde analisado limita a generalização dos resultados. Esse tipo de análise, portanto, não forneceu *insights* consistentes sobre o comportamento da ferramenta em relação a padrões de substituição ou simplificação harmônica.

7 DISCUSSÃO

A avaliação do sistema revelou diversos pontos positivos, mas também expôs limitações que devem ser consideradas ao interpretar os resultados. Embora as métricas analisadas forneçam uma visão geral do desempenho da ferramenta na extração e formatação de cifras, existem fatores que podem ter influenciado os resultados obtidos. A seguir, são apresentadas algumas das principais limitações identificadas durante o processo de validação.

7.1 SIMPLIFICAÇÃO DOS ACORDES

Para as análises de posicionamento e de sequências, foram utilizadas formas simplificadas dos acordes, considerando apenas suas raízes. Essa abordagem foi adotada para minimizar o impacto de variações na nomenclatura (como inversões e extensões) e para identificar casos em que a ferramenta pode ter simplificado acordes mais complexos. No entanto, essa simplificação também pode ter levado a uma subestimação de certos erros, já que acordes diferentes podem ter sido tratados como equivalentes.

7.2 CRITÉRIO DE POSICIONAMENTO DOS ACORDES

A avaliação do posicionamento dos acordes levou em conta um desvio de até ± 2 posições, considerando que pequenas variações podem ocorrer devido a espaçamentos irregulares ou diferenças na disposição dos acordes sobre a letra. Esse critério buscou tornar a análise mais tolerante a pequenas imprecisões na formatação. No entanto, em alguns casos, esse desvio pode ter sido insuficiente para captar erros mais sutis ou excessivamente permissivo para certas cifras, afetando a precisão da métrica de alinhamento.

7.3 CRITÉRIO DE CORRESPONDÊNCIA DOS ACORDES

A comparação entre os acordes gerados e os da cifra de validação considerou apenas a presença ou ausência de cada acorde, sem levar em conta sua função harmônica no contexto musical. Em alguns casos, acordes podem ser substituídos por variações equivalentes sem comprometer a sonoridade da música, como a troca de um acorde maior por sua sétima ou de uma inversão por sua forma fundamental. Assim, a métrica de acordes em comum pode não

refletir com precisão a similaridade real entre as cifras, já que algumas substituições podem ser musicalmente aceitáveis.

7.4 VARIAÇÃO ENTRE VERSÕES DA MESMA MÚSICA

Sempre que possível, foram utilizados os vídeos disponibilizados pela ferramenta de cifras para garantir a correspondência entre áudio e cifra. No entanto, em alguns casos, esses vídeos não estavam disponíveis para download, tornando necessário recorrer a outras versões da música. Isso pode ter introduzido pequenas diferenças nos acordes e na estrutura da cifra, uma vez que versões ao vivo, acústicas ou regravações podem apresentar mudanças nos arranjos harmônicos. Assim, parte da divergência entre as cifras geradas e as cifras de validação pode estar relacionada a essas variações naturais entre diferentes interpretações da mesma música.

7.5 ANÁLISE QUANTITATIVA VS. QUALITATIVA

As métricas utilizadas se baseiam em comparações quantitativas, o que significa que a análise foi fundamentada na presença, frequência e posição dos acordes. No entanto, a avaliação qualitativa da cifra gerada, como a percepção auditiva da harmonia resultante, não foi considerada diretamente.

7.6 DEPENDÊNCIA DA CIFRA DE VALIDAÇÃO

Além disso, a dependência da cifra de validação como referência pode introduzir viés na análise, pois eventuais erros ou variações nessa cifra podem impactar os resultados. Algumas cifras de validação podem não representar a versão definitiva ou mais precisa da música, o que pode ter levado a discrepâncias na comparação.

8 CONCLUSÃO

A automatização da formatação de cifras representa um avanço significativo no campo da educação musical e na prática instrumental. Ao facilitar a transcrição e o alinhamento de acordes com as letras das músicas, essa tecnologia tem o potencial de transformar a maneira como músicos interagem com o repertório musical.

Para os músicos, especialmente aqueles em fase de aprendizado, ferramentas automatizadas podem servir como auxiliares valiosos, permitindo uma compreensão mais rápida e precisa das progressões harmônicas. Isso não apenas economiza tempo, mas também melhora a precisão na execução musical. Além disso, ao reduzir a dependência de transcrições manuais, que podem ser demoradas e sujeitas a erros, os músicos podem se concentrar mais na interpretação e expressão artística.

Na esfera educacional, a integração de tecnologias de transcrição automática pode enriquecer o processo de ensino e aprendizagem. Educadores podem utilizar essas ferramentas para demonstrar conceitos harmônicos de forma mais interativa, proporcionando aos alunos uma experiência prática e imediata. Além disso, a disponibilidade de cifras precisas e bem formatadas pode incentivar a prática autônoma dos estudantes, promovendo um aprendizado mais eficaz e engajador.

No entanto, é importante reconhecer que, apesar dos avanços, desafios persistem na precisão da extração e posicionamento dos acordes. Erros como a omissão de acordes importantes ou a inserção excessiva de acordes comuns podem comprometer a fidelidade das cifras geradas. Portanto, esforços contínuos são necessários para aprimorar a tecnologia, visando atender às necessidades dos músicos e educadores com maior eficácia.

8.1 DIREÇÕES PARA MELHORIAS FUTURAS

A análise dos resultados sugere várias áreas de aprimoramento para aumentar a precisão da ferramenta:

- Melhoria na Detecção de Acordes em Trechos Cantados: Foi observado que os acordes são melhores identificados em partes instrumentais, enquanto nas seções cantadas algumas notas deixam de ser reconhecidas ou aparecem de forma simplificada (por exemplo, E em vez de E7). Técnicas de separação de fontes podem ser exploradas para isolar a

faixa instrumental e melhorar a detecção harmônica.

- Inclusão de Acordes em Partes Estruturalmente Importantes: No modelo atual, apenas os acordes diretamente associados às partes cantadas são considerados na formatação final, o que pode levar à omissão de harmonias presentes em introduções, *bridges* e seções instrumentais. Uma abordagem futura pode envolver a análise dessas seções separadamente para entender seu impacto na estrutura harmônica da música, podendo ser incorporadas na cifra final.

- Aprimoramento do Posicionamento dos Acordes: Melhorias na exibição dos acordes sobre a letra podem ser feitas para reduzir desalinhamentos. Métodos mais robustos de sincronização entre harmonia e melodia podem contribuir para um alinhamento mais preciso.

- Exploração de Diferentes Ferramentas para Detecção de Acordes: Considerando que a ferramenta atual pode ter limitações para músicas cantadas, pode ser interessante avaliar outras abordagens ou modelos especializados nesse tipo de análise. Além disso, parece haver inconsistências na identificação do momento exato em que os acordes ocorrem dentro da música. O uso de ferramentas mais especializadas para detecção temporal pode ajudar a melhorar a precisão tanto na extração dos acordes quanto no seu posicionamento na cifra.

- Uso das Letras das Próprias Cifras para Validação: Atualmente, as letras e as cifras podem ser extraídas de fontes diferentes, o que pode gerar inconsistências na comparação, especialmente na análise do posicionamento dos acordes. Utilizar as letras presentes nas próprias cifras para a validação permitiria uma correspondência mais precisa, reduzindo o impacto de variações textuais e possibilitando uma avaliação mais robusta do alinhamento dos acordes em relação às palavras da música.

- Uso de Cifras Verificadas para Validação: Para aumentar a confiabilidade dos dados de validação, uma possível melhoria seria utilizar apenas cifras verificadas na plataforma de extração. Isso reduziria inconsistências nas comparações e garantiria uma base mais sólida para avaliação.

REFERÊNCIAS

CHEDIAK, A. *Harmonia & Improvisaçao - Vol. 1*. LUMIAR, 1986. (Harmonia & improvisaçao: 70 músicas harmonizadas e analisadas : viol˜ao, guitarra, baixo, teclado). ISBN 9788585426064. Disponível em: <https://books.google.com.br/books?id=LU3PhggS1AMC>.

CONTRIBUTORS yt-dlp. *yt-dlp: A feature-rich command-line audio/video downloader*. 2025. Acesso em: 1 abr. 2025. Disponível em: <https://github.com/yt-dlp/yt-dlp>.

DALMAZZO, D. C.; DÉGUERNEL, K.; STURM, B. L. T. The chordinator: Chord progression modeling and generation using transformers. In: *Proceedings of the 24th International Society for Music Information Retrieval Conference (ISMIR 2023)*. [s.n.], 2023. Disponível em: <https://ismir2023program.ismir.net/lbd_316.html>.

DOBBERT, F. *Como Funciona o Sistema de Cifras*. 2020. <https://blog.fritzdobbert.com.br/tudo-sobre-piano/como-funciona-sistema-de-cifras/>. Acessado em: 10 jul. 2024.

EDWARDS, A. C.; RILEY, X.; SARMENTO, P. P.; DIXON, S. Midi-to-tab: Guitar tablature inference via masked language modeling. In: *Proceedings of the 25th International Society for Music Information Retrieval Conference*. ISMIR, 2024. p. 288–294. Disponível em: <https://doi.org/10.5281/zenodo.14877333>.

FOUNDATION, P. S. *collections — Tipos de dados contêineres especializados*. 2025. Acesso em: 1 abr. 2025. Disponível em: <https://docs.python.org/3/library/collections.html>.

FOUNDATION, P. S. *difflib — Helpers for computing deltas*. 2025. Acesso em: 1 abr. 2025. Disponível em: <https://docs.python.org/3/library/difflib.html>.

FOUNDATION, P. S. *Python 3.9.5*. 2025. Acesso em: 1 abr. 2025. Disponível em: <https://www.python.org/>.

GAO, X.; GUPTA, C.; LI, H. Automatic lyrics transcription of polyphonic music with lyrics-chord multi-task learning. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 30, p. 2280–2294, 2022.

Genius API. *Genius API*. 2025. Acesso em: 01 abr. 2025. Disponível em: <https://genius.com/developers>.

GITHUB, I. *GitHub*. 2025. Acesso em: 1 abr. 2025. Disponível em: <https://github.com/>.

GOOGLE, I. *Google Colab*. 2025. Acesso em: 1 abr. 2025. Disponível em: <https://colab.research.google.com/>.

INC., M. S. *Moises: The Musician's App*. 2025. Accessed: 2025-04-04. Disponível em: <https://moises.ai/>.

JIANG, J.; CHEN, K.; LI, W.; XIA, G. G. Large-vocabulary chord transcription via chord structure decomposition. In: *International Society for Music Information Retrieval Conference*. [s.n.], 2019. Disponível em: <https://api.semanticscholar.org/CorpusID:208334209>.

MCVICAR, M.; SANTOS-RODRíGUEZ, R.; NI, Y.; BIE, T. D. Automatic chord estimation from audio: A review of the state of the art. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 22, n. 2, p. 556–575, 2014.

MICCHI, G.; KOSTA, K.; MEDEOT, G.; CHANQUION, P. A deep learning method for enforcing coherence in automatic chord recognition. In: *Proceedings of the 22nd International Society for Music Information Retrieval Conference*. ISMIR, 2021. p. 443–451. Disponível em: <https://doi.org/10.5281/zenodo.5624539>.

MIHIRA, Y. *pychord: Biblioteca Python para manipulação de acordes musicais*. 2023. Acesso em: 1 abr. 2025. Disponível em: <https://github.com/yuma-m/pychord>.

ODEKERKEN, D.; KOOPS, H. V.; VOLK, A. Improving audio chord estimation by alignment and integration of crowd-sourced symbolic music. *Transactions of the International Society for Music Information Retrieval*, v. 4, n. 1, p. 141–155, 2021. Disponível em: <https://transactions.ismir.net/articles/10.5334/tismir.81/>.

OLIPHANT, T. *NumPy: Pacote fundamental para computação científica com Python*. 2024. Acesso em: 1 abr. 2025. Disponível em: <https://numpy.org/>.

OPENAI. *Whisper: Automatic Speech Recognition*. 2022. <https://github.com/openai/whisper>. Acesso em: [data de acesso].

OU, N. A. *Chord ai: Identificação Automática de Acordes e Batidas*. 2025. Disponível para dispositivos iOS e Android. Disponível em: <https://chordai.net/>.

POLTRONIERI, A.; PRESUTTI, V.; ROCAMORA, M. Chordsync: Conformer-based alignment of chord annotations to music audio. *ArXiv*, abs/2408.00674, 2024. Disponível em: <https://api.semanticscholar.org/CorpusID:271297123>.

READBEYOND. *Aeneas: A lightweight forced aligner*. <https://github.com/readbeyond/aeneas>. Acesso em: [data de acesso].

REITZ, K. *requests: HTTP for Humans*. 2024. Acesso em: 1 abr. 2025. Disponível em: <https://requests.readthedocs.io/>.

RICHARDSON, L. *Beautiful Soup Documentation*. 2015. Acesso em: 1 abr. 2025. Disponível em: <https://beautiful-soup-4.readthedocs.io/en/latest/>.

ROUSSET, F. *Demucs: Deep Extractor for Music Sources*. 2021. Acesso em: 01 abr. 2025. Disponível em: <https://github.com/facebookresearch/demucs>.

ROWE, L.; TZANETAKIS, G. Curriculum learning for imbalanced classification in large vocabulary automatic chord recognition. In: *Proceedings of the 22nd International Society for Music Information Retrieval Conference (ISMIR)*. [s.n.], 2021. p. 1–8. Disponível em: <https://zenodo.org/record/5624463>.

SCHEDL, M.; GóMEZ, E.; URBANO, J. Music information retrieval: Recent developments and applications. *Foundations and Trends in Information Retrieval*, v. 8, p. 127–261, 01 2014.

SOL, S. *Cifra Club - Cifras e Tablaturas de Músicas*. 2025. Acesso em: 1 abr. 2025. Disponível em: <https://www.cifraclub.com.br>.

SOL, S. *Letras.mus.br - Letras de Músicas*. 2025. Acesso em: 1 abr. 2025. Disponível em: <https://www.letras.mus.br>.

TABORDA, M. *Violão e identidade nacional: Rio de Janeiro, 1830-1930*. Civilização Brasileira, 2011. ISBN 9788520010297. Disponível em: <https://books.google.com.br/books?id=zwKSZwEACAAJ>.

YOUTUBE, I. *YouTube*. 2025. Acesso em: 1 abr. 2025. Disponível em: <https://www.youtube.com/>.

ZHOU, M. *lyrics-sync: Sistema de alinhamento de áudio e letras baseado em aprendizado profundo*. 2023. Acesso em: 1 abr. 2025. Disponível em: <https://github.com/mikezzb/lyrics-sync>.

APÊNDICE A – DICIONÁRIO DE INVERSÕES DE ACORDES

Inversão	Acordes Possíveis
C/3	C/E, C6/E, Cadd9/E
D/3	D/F#, D6/F#, D7/F#
E/3	E/G#, E6/G#, Eadd9/G#
F/3	F/A, F6/A, F7/A
G/3	G/B, G6/B, G7/B, G4/B
A/3	A/C#, A6/C#, A7/C#
B/3	B/D#, B6/D#, B7/D#
Cm/3	Cm/Eb, Cm6/Eb, Cm7/Eb
Dm/3	Dm/F, Dm6/F, Dm7/F
Em/3	Em/G, Em6/G, Em7/G
Fm/3	Fm/Ab, Fm6/Ab, Fm7/Ab
Gm/3	Gm/Bb, Gm6/Bb, Gm7/Bb
Am/3	Am/C, Am6/C, Am7/C
Bm/3	Bm/D, Bm6/D, Bm7/D
C/5	C/G, C7/G, C6/G
D/5	D/A, D7/A, D6/A
E/5	E/B, E7/B, E6/B
F/5	F/C, F7/C, F6/C
G/5	G/D, G7/D, G6/D
A/5	A/E, A7/E, A6/E
B/5	B/F#, B7/F#, B6/F#
Cm/5	Cm/G, Cm7/G, Cm6/G
Dm/5	Dm/A, Dm7/A, Dm6/A

Continua na próxima página

Inversão	**Acordes Possíveis**
Em/5	Em/B, Em7/B, Em6/B
Fm/5	Fm/C, Fm7/C, Fm6/C
Gm/5	Gm/D, Gm7/D, Gm6/D
Am/5	Am/E, Am7/E, Am6/E
Bm/5	Bm/F#, Bm7/F#, Bm6/F#
C7/3	C7/E, C9/E
D7/3	D7/F#, D9/F#
E7/3	E7/G#, E9/G#
F7/3	F7/A, F9/A
G7/3	G7/B, G9/B
A7/3	A7/C#, A9/C#
B7/3	B7/D#, B9/D#
Cm7/3	Cm7/Eb, Cm9/Eb
Dm7/3	Dm7/F, Dm9/F
Em7/3	Em7/G, Em9/G
Fm7/3	Fm7/Ab, Fm9/Ab
Gm7/3	Gm7/Bb, Gm9/Bb
Am7/3	Am7/C, Am9/C
Bm7/3	Bm7/D, Bm9/D
C7/5	C7/G, C9/G
D7/5	D7/A, D9/A
E7/5	E7/B, E9/B
F7/5	F7/C, F9/C
G7/5	G7/D, G9/D
A7/5	A7/E, A9/E

Continua na próxima página

Inversão	**Acordes Possíveis**
B7/5	B7/F#, B9/F#
Cm7/5	Cm7/G, Cm9/G
Dm7/5	Dm7/A, Dm9/A
Em7/5	Em7/B, Em9/B
Fm7/5	Fm7/C, Fm9/C
Gm7/5	Gm7/D, Gm9/D
Am7/5	Am7/E, Am9/E
Bm7/5	Bm7/F#, Bm9/F#

Tabela 2 – Inversões de acordes e suas variantes

APÊNDICE B – BASE DE DADOS

Música - Áudio - Letra - Cifra
Que País É Este Áudio: <https://youtu.be/kFpkyT_KPpc> Letra: <https://www.letras.mus.br/legiao-urbana/46973/> Cifra: <https://www.cifraclub.com.br/legiao-urbana/que-pais-e-esse/>
Infiel Áudio: <https://youtu.be/eCyMh-mZ1B0> Letra: <https://www.letras.mus.br/marilia-mendonca/infiel/> Cifra: <https://www.cifraclub.com.br/marilia-mendonca/infiel/>
Tocando Em Frente Áudio: <https://youtu.be/CugYXgJ2SFI?t=11> Letra: <https://www.letras.mus.br/almir-sater/44082/> Cifra: <https://www.cifraclub.com.br/almir-sater/tocando-em-frente/>
Petrolina Juazeiro Áudio: <https://youtu.be/VchbqjKk6wA> Letra: <https://www.letras.mus.br/alceu-valenca/400650/> Cifra: <https://www.cifraclub.com.br/alceu-valenca/petrolina-juazeiro/>
Naquela Mesa Áudio: <https://youtu.be/KkCre5ZEbVU> Letra: <http://letras.mus.br/nelson-goncalves/47663/> Cifra: <https://www.cifraclub.com.br/nelson-goncalves/naquela-mesa/>
Uma Carta Áudio: <https://youtu.be/ePFQIiLI1G8> Letra: <https://www.letras.mus.br/ls-jack/72378/> Cifra: <https://www.cifraclub.com.br/ls-jack/uma-carta/>

Música - Áudio - Letra - Cifra
Como Vai Você Áudio: <https://youtu.be/dHEqjSdDKok> Letra: <https://www.letras.mus.br/roberto-carlos/48571/> Cifra: <https://www.cifraclub.com.br/roberto-carlos/como-vai-voce/>
Evidências Áudio: <https://youtu.be/Y59pC4FcBxM> Letra: <https://www.letras.mus.br/chitaozinho-e-xororo/768469/> Cifra: <https://www.cifraclub.com.br/chitaozinho-e-xororo/evidencias/>
Sozinho Áudio: <https://youtu.be/lnjZ2pqI6Z8> Letra: <https://www.letras.mus.br/caetano-veloso/41672/> Cifra: <https://www.cifraclub.com.br/caetano-veloso/sozinho/>
Garota de Ipanema Áudio: <https://youtu.be/WuenyQ4NCQE> Letra: <https://www.letras.mus.br/tom-jobim/20018/> Cifra: <https://www.cifraclub.com.br/tom-jobim/garota-de-ipanema/>
João e Maria Áudio: <https://youtu.be/wqiKeSfdsrc> Letra: <https://www.letras.mus.br/chico-buarque/45140/> Cifra: <https://www.cifraclub.com.br/chico-buarque/joao-maria/>
Águas de Março Áudio: <https://youtu.be/g2P_OupxkB8> Letra: <https://www.letras.mus.br/tom-jobim/49022/> Cifra: <https://www.cifraclub.com.br/tom-jobim/aguas-de-marco/>
Corcovado Áudio: <https://youtu.be/25nv3865Ak0> Letra: <https://www.letras.mus.br/tom-jobim/49031/> Cifra: <https://www.cifraclub.com.br/tom-jobim/corcovado/>

Música - Áudio - Letra - Cifra
Summertime Sadness Áudio: <https://youtu.be/nVjsGKrE6E8> Letra: <https://www.letras.mus.br/lana-del-rey/1999476/> Cifra: <https://www.cifraclub.com.br/lana-del-rey/summertime-sadness/>
Beautiful Things Áudio: <https://youtu.be/Oa_RSwwpPaA> Letra: <https://www.letras.mus.br/benson-boone/beautiful-things/> Cifra: <https://www.cifraclub.com.br/benson-boone/beautiful-things/>
It's My Life Áudio: <https://youtu.be/vx2u5uUu3DE> Letra: <https://www.letras.mus.br/bon-jovi/4924/> Cifra: <https://www.cifraclub.com.br/bon-jovi/its-my-life/>
Always On My Mind Áudio: <https://youtu.be/u9sRJ-eOHnc> Letra: <https://www.letras.mus.br/elvis-presley/31393/> Cifra: <https://www.cifraclub.com.br/elvis-presley/always-on-my-mind/>
Love Me Tender Áudio: <https://youtu.be/YC0BPcLea0Y> Letra: <https://www.letras.mus.br/elvis-presley/31460/> Cifra: <https://www.cifraclub.com.br/elvis-presley/love-me-tender/>
My Sacrifice Áudio: <https://youtu.be/O-fyNgHdmLI> Letra: <https://www.letras.mus.br/creed/9093/> Cifra: <https://www.cifraclub.com.br/creed/my-sacrifice/>
Wherever You Will Go / Fotos Áudio: <https://youtu.be/Zi8mrT4l6vg> Letra: <https://www.letras.mus.br/lu-e-robertinho/wherever-you-will-go-fotos/> Cifra: <https://www.cifraclub.com.br/lu-e-robertinho/wherever-you-will-go--fotos/>

Tabela 3 – Lista de músicas com URLs de áudio, letra e cifra.