

HeroKeys: Um ambiente virtual gamificado para tocar e aprender piano

Deivison R. Jordão
UFPB
João Pessoa, Brasil
drj@academico.ufpb.br

Joaquim B. Cavalcante
UFPB
João Pessoa, Brasil
joaquim.breno@academico.ufpb.br

Humberto Navarro de Carvalho
Music.AI
João Pessoa, Brasil
humberto.navarro@moises.ai

Yvson Nunes
Music.AI
João Pessoa, Brasil
yvson.nunes@moises.ai

Thais Gaudencio
ARIA/UFPB
João Pessoa, Brasil
gaudenciothais@gmail.com

Yuri Malheiros
ARIA/UFPB
João Pessoa, Brasil
yuri@ci.ufpb.br

ABSTRACT

Musical education remains inaccessible to many due to the reliance on traditional methods, which demand sheet music, prior knowledge, or formal training. This paper presents HeroKeys, an open-source gamified web application designed to enhance piano learning through artificial intelligence-based audio transcription. Unlike traditional learning methods that rely on pre-existing transcriptions, HeroKeys allows users to practice any song of their choice by automatically, converting audio files into playable musical notation. The system features multiple visualization options including *Piano Roll*, traditional sheet music, and a display of the current chord being played, accommodating different learning preferences. The application incorporates gamification elements such as scoring systems, multipliers, and visual feedback to increase user engagement and motivation. HeroKeys supports external MIDI keyboard connections and provides customizable playback controls to facilitate progressive learning. Our approach combines audio processing technologies with interactive elements to create an accessible and motivating piano learning environment, making musical education more engaging and accessible to beginners.

KEYWORDS

Aprendizagem de piano, transcrição musical, inteligência artificial, gamificação

1 INTRODUÇÃO

Para aspirantes a músicos, poder tocar suas músicas favoritas é um dos maiores estímulos para seguir aprendendo. No contexto do piano, existem diversas abordagens para aprender uma nova peça musical: treino auditivo, videoaulas, leitura de partituras, tablaturas e cifras, ou interfaces interativas que apresentam as notas em tempo real. Cada método possui suas vantagens, entretanto eles enfrentam limitações práticas: nem todas as músicas possuem transcrições disponíveis, e o treino auditivo exige habilidades musicais que muitos iniciantes ainda não desenvolveram.

Nesse âmbito, este trabalho apresenta o HeroKeys, uma aplicação *open-source* gamificada para aprendizado de piano que utiliza aprendizagem de máquina para transcrever automaticamente arquivos de áudio. A ferramenta isola a faixa de piano através de técnicas de separação de fontes [2] e, em seguida, aplica métodos de transcrição musical automática [7] para converter o áudio isolado em arquivo MIDI, permitindo que o usuário pratique qualquer música de sua escolha. Ao possibilitar a transcrição de qualquer faixa escolhida pelo próprio usuário, o HeroKeys elimina a dependência de materiais previamente transcritos e personaliza o repertório de estudo, tornando a prática mais imersiva. As notas transcritas podem ser exibidas na tela como notas individuais ou em partitura tradicional. O sistema também oferece suporte à utilização de um teclado externo MIDI, no qual o usuário pode conectar o teclado por USB e tocar as notas correspondentes à música. Como elemento gamificado, a aplicação fornece uma resposta visual em tempo real, reportando através da interface o desempenho do usuário a cada nota executada.

O artigo está estruturado em seis seções: a Seção 2 revisa trabalhos relacionados; a Seção 3 detalha a metodologia, incluindo arquitetura, transcrição automática, visualização interativa, gamificação e análise comparativa; a Seção 4 aborda a conectividade e o controle via interface MIDI; a Seção 5 apresenta um guia de uso do HeroKeys; e a Seção 6 traz as conclusões e perspectivas de evolução da ferramenta.

2 TRABALHOS RELACIONADOS

Diversas abordagens têm sido desenvolvidas para facilitar o aprendizado de piano, combinando tecnologias de visualização, transcrição e elementos de gamificação [6, 8, 11]. Ferramentas online como piano2notes [6] oferecem funcionalidades de transcrição que convertem gravações de piano em notação musical. Embora úteis para músicos que buscam transcrever performances, estas ferramentas geralmente apresentam limitações na experiência de aprendizado, focando apenas na transcrição sem elementos que engajem ativamente o usuário no processo educacional.

No trabalho [11] é explorado o uso de processamento de áudio em ambientes de realidade mista para o aprendizado de piano. Utilizando o HoloLens 2 e o método *Fast Fourier transform* (FFT) com detecção de picos, os autores alcançaram 97,28 % de precisão na identificação de notas em músicas complexas. Seu estudo revelou que

In: XXIV Workshop de Ferramentas e Aplicações (WFA 2025). Anais Estendidos do XXXI Simpósio Brasileiro de Sistemas Multimídia e Web (WFA'2025). Rio de Janeiro/RJ, Brasil. Porto Alegre: Brazilian Computer Society, 2025.
© 2025 SBC – Sociedade Brasileira de Computação.
ISSN 2596-1683

curiosidade e diversão foram os principais fatores motivacionais, resultando em um aumento de 31,28% na eficácia do aprendizado.

De forma similar, no trabalho [8] foi desenvolvido o "Mixed Reality Piano Tutor", um ambiente gamificado que utiliza realidade mista para reduzir a carga cognitiva durante o aprendizado. O sistema inclui um *Piano roll* 3D e *feedback* visual em tempo real sobre a precisão das notas, demonstrando maior eficácia na execução de notas, quando comparado a sistemas comerciais tradicionais.

Diferente das soluções de realidade mista [8, 11] que exigem equipamentos especializados, nossa aplicação funciona em dispositivos convencionais. Diferentemente da maioria das soluções existentes, que dependem de arquivos MIDI, o sistema proposto é capaz de processar diretamente arquivos de áudio, inclusive em gravações polifônicas contendo múltiplos instrumentos, desde que haja a presença do piano. Essa abordagem amplia de forma significativa o repertório disponível para prática, tornando o aprendizado musical mais acessível e motivador.

3 METODOLOGIA

A aplicação foi arquitetada com um *frontend* em Next.js [10], que carrega o áudio de entrada e realiza duas operações principais. Primeiro, o áudio é enviado para o módulo de separação de piano da API do Music.ai [9], a fim de isolar a faixa de piano. Em seguida, o *frontend* utiliza o mesmo áudio para aplicar o módulo de reconhecimento de acordes do Music.ai, identificando os acordes presentes na música. Após a separação da faixa de piano, o áudio isolado é enviado ao *backend* implementado em FastAPI [4]. O *backend* transforma esse áudio de piano em um arquivo MIDI utilizando o modelo de transcrição de piano de alta resolução com pedais [7]. A partir do arquivo MIDI e dos dados de acordes, o *frontend* exibe visualizações como acordes detectados, notas em *Piano Roll* ou em partitura tradicional, conforme ilustrado na Figura 1. O código-fonte do HeroKeys é disponibilizado como software livre através do GitHub sob a licença MIT¹.

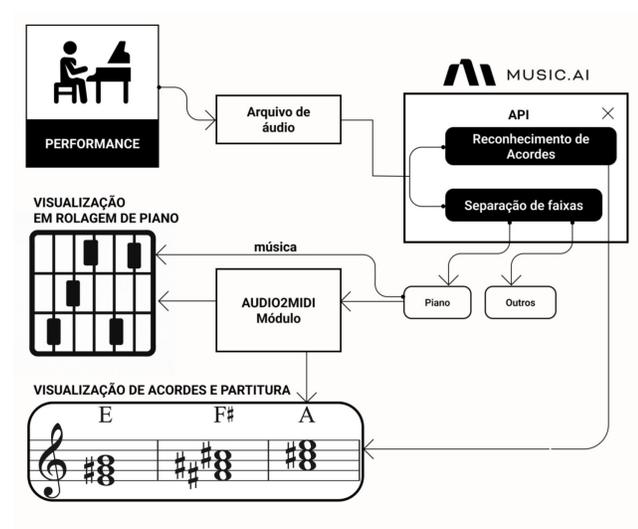


Figure 1: Fluxograma de processamento do HeroKeys.

¹<https://opensource.org/license/mit>

3.1 CARREGAMENTO E PROCESSAMENTO

O carregamento dos dados é gerenciado por uma interface apresentada na Figura 2, que foi desenvolvida para ser intuitiva, permitindo tanto o método de arrastar e soltar, quanto a seleção tradicional de arquivos. O sistema é compatível com os mesmos formatos de áudio disponibilizados pelo Music.ai, abrangendo os principais padrões utilizados no mercado e garantindo ampla integração com diversas fontes sonoras.

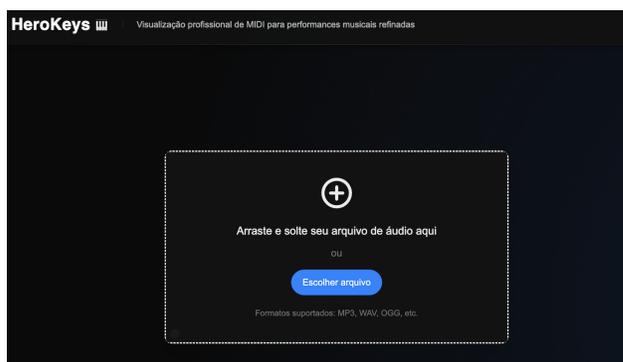


Figure 2: Página Inicial.

Após o carregamento, o arquivo passa por um *pipeline* de processamento:

- **Reconhecimento de Acordes:** Paralelamente, o módulo de reconhecimento de acordes do Music.AI [9] analisa o arquivo para identificar a progressão harmônica.
- **Separação de Fontes:** Utilizando os módulos especializados em decomposição de áudio através de redes neurais profundas do Music.AI, o sistema isola a faixa de piano dos demais elementos sonoros.
- **Transcrição de Notas:**

A faixa de piano isolada é analisada por um modelo de transcrição de áudio para MIDI de alta resolução [7], que regrida os tempos precisos de início e término das notas. O sistema funciona significativamente melhor com faixas isoladas de piano, já que foi especificamente treinado e desenvolvido para transcrição de piano e não para transcrição multi-instrumental.

O sistema utiliza uma arquitetura híbrida onde redes neurais convolucionais (*Convolutional Neural Networks* - CNNs) são empregadas para extração hierárquica de características espectrais do áudio. As CNNs processam espectrogramas log-mel através de múltiplos blocos convolucionais com *pooling* aplicado exclusivamente no eixo de frequência, preservando a resolução temporal necessária para detecção precisa de *onsets* e *offsets*. Cada bloco convolucional utiliza filtros com *kernels* 3×3 para capturar padrões locais de ataques de notas e estruturas harmônicas, seguidos por normalização em lote e ativação ReLU. As características extraídas pelas CNNs são então processadas por unidades bidirecionais de portões recorrentes (*bidirectional Gated Recurrent Units* - biGRUs), que modelam dependências temporais de longo prazo e alimentam quatro submódulos especializados: regressão de

velocidade, regressão de início de notas (*onset*), classificação de quadros e regressão de término de notas (*offset*). Diferente do Basic Pitch, que utiliza o formato *Constant-Q Transform - CQT*, o sistema especializado em piano processa espectrogramas log-mel e alcança medida F1 de 91,86% na detecção de notas, e eventos de pedal de sustentação, demonstrando como a arquitetura CNN otimizada para características espectrais do piano resulta em refinamento progressivo da detecção e impacto significativo na experiência do usuário.

3.2 VISUALIZAÇÃO

O HeroKeys oferece múltiplas formas de visualização dos elementos musicais, atendendo a diferentes níveis de conhecimento e preferências dos usuários.

A visualização principal do sistema é o *Piano Roll*, como pode ser observado na Figura 3. Nela, as notas são exibidas como blocos retangulares em uma grade bidimensional, onde o eixo vertical representa o tempo e o eixo horizontal representa a altura das notas. As notas são codificadas por cores para indicar seu estado, se estão por vir ou ativas, e descem verticalmente em perfeita sincronização com o áudio. Na parte inferior, o teclado do piano é mostrado e se ilumina em tempo real conforme as notas são tocadas.



Figure 3: *Piano Roll* e acordes.

Na parte superior da interface, há uma visualização específica para acordes, também ilustrada nas Figuras 3 e 4. Os acordes aparecem em uma sequência horizontal com nomenclatura musical padrão. Durante a reprodução, o acorde atual é destacado visualmente e são mostrados seu nome e o tempo exato de início.

A aplicação oferece a opção de visualização tradicional através de partitura com notação padrão em pentagramas com claves de sol e fá, como exposto na Figura 4. A partitura avança automaticamente em tempo real durante a reprodução e apresenta informações de ritmo, dinâmica e acidentes. O *layout* se adapta de forma responsiva, ajustando tamanho e densidade para garantir legibilidade.

3.3 GAMIFICAÇÃO

O HeroKeys incorpora elementos de gamificação inspirados em jogos musicais como Guitar Hero [5], como bem ilustrado na Figura 3, fundamentando-se na Teoria da Autodeterminação (autonomia via escolha de músicas, competência via *feedback* imediato) e na Teoria do Flow (equilíbrio desafio-habilidade através de controles de velocidade e segmentação) [1, 3]. Cada nota executada corretamente

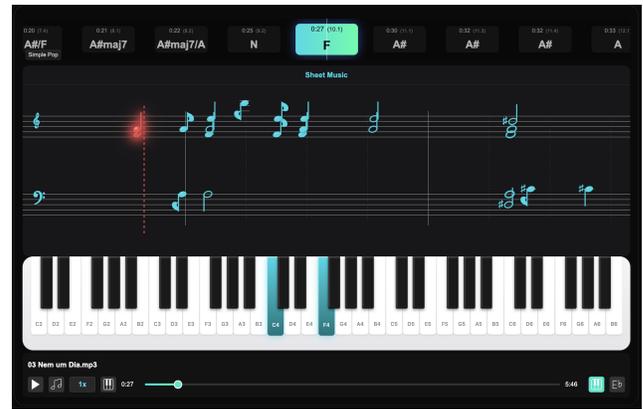


Figure 4: Visualização em Partitura.

rende pontos de acordo com a precisão temporal, com *feedback* visual imediato a cada ação e exibição proeminente da pontuação total acumulada na interface. Para reforçar a consistência, o sistema acompanha a sequência de acertos (*streak*) e aplica multiplicadores progressivos de pontuação (x1,0, x1,5 e x2,0), que aumentam conforme o usuário mantém acertos consecutivos, sendo redefinidos quando uma nota é tocada incorretamente ou deixada de tocar. A implementação desses sistemas de pontuação, multiplicadores e *streak* alinha-se aos "sistemas de pontuação e recompensas intrínsecas" validados como produtores de "impactos significativos" no engajamento educacional [12]. A experiência é enriquecida por elementos visuais dinâmicos, incluindo efeitos de partículas coloridas que celebram pontuações alcançadas e animações de interface que reagem às ações do usuário, constituindo uma convergência teórica que legitima os elementos gamificados como métodos pedagogicamente fundamentados para motivação intrínseca no ensino de piano, transcendendo o mero entretenimento.

3.4 CONECTIVIDADE E CONTROLE

O HeroKeys implementa uma interface MIDI que permite a conexão de teclados e controladores MIDI externos via USB. Os dispositivos conectados são detectados automaticamente, com mapeamento das notas para a aplicação e captura da atividade para avaliar a dinâmica da performance.

Além disso, o HeroKeys oferece controles de reprodução e áudio que personalizam a prática: é possível ajustar a velocidade de reprodução para 0,25x, 0,50x, 1x, 1,25x ou 1,5x; ativar ou silenciar o áudio original da música ou, inversamente, silenciar o áudio do teclado do usuário. A navegação é realizada por controles de reprodução/pausa e por uma linha do tempo interativa que permite acesso rápido a qualquer ponto da música.

4 USANDO HEROKEYS

Ao inicializar a aplicação, o usuário encontra a interface inicial representada na Figura 2 onde pode inserir seu arquivo de áudio na área de *upload*. O sistema processa automaticamente o material, separando a faixa de piano, transcrevendo as notas para formato MIDI e identificando a progressão harmônica.

Paralelamente, o usuário pode conectar um teclado MIDI via USB, que é detectado automaticamente pelo HeroKeys com confirmação visual da conexão.

Concluído o processamento, a interface apresenta a *Piano Roll* exposto na Figura 3 com notas musicais visualizadas como elementos retangulares descendentes. O usuário tem a opção de alternar para a visualização em Partitura tradicional, representado Figura 4, conforme sua preferência de leitura musical. Para peças de maior complexidade, o controle de velocidade permite ajustar o andamento, facilitando o aprendizado gradual.

Durante a execução, o sistema de pontuação fornece ao usuário *feedback* em tempo real sobre a qualidade da performance. A navegação é simplificada pela linha temporal interativa, permitindo que o usuário retorne ou avance para segmentos específicos da música, otimizando o processo de prática em trechos que exigem maior atenção.

5 ANÁLISE COMPARATIVA

A Tabela 1 apresenta uma comparação entre o HeroKeys e sistemas existentes para aprendizado de piano. Os sistemas de Realidade Mista [8, 11] oferecem *feedback* em tempo real e gamificação, mas requerem equipamentos especializados. O piano2notes [6] proporciona transcrição automática em dispositivos convencionais, porém carece de interatividade. O HeroKeys combina acessibilidade, transcrição automática, *feedback* em tempo real e gamificação, incorporando ainda visualização múltipla, reconhecimento de acordes e suporte a qualquer arquivo de áudio, configurando-se como uma alternativa completa e acessível ao aprendizado musical.

Table 1: Comparação de Sistemas de Aprendizado de Piano

Recursos	RM ¹ [8, 11]	p2n ² [6]	HeroKeys
Equip. Acessível	×	✓	✓
Transcrição Automática	×	✓	✓
Feedback Real-time	✓	×	✓
Gamificação	✓	×	✓
Visualização Múltipla	×	×	✓
Reconhecimento Acordes	×	×	✓
Suporte Qualquer Áudio	×	×	✓

¹RM: Sistemas de Realidade Mista

²p2n: piano2notes

6 CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho apresentou o HeroKeys², uma aplicação *open-source*³ para aprendizado de piano que integra transcrição automática de áudio por inteligência artificial (IA) com elementos de gamificação. A ferramenta permite que usuários pratiquem músicas de sua escolha sem depender de transcrições pré-existentes, constituindo uma alternativa inovadora aos métodos tradicionais de ensino musical. Por meio da visualização de notas e resposta em tempo real, combinadas com suporte a teclados MIDI via USB, a aplicação oferece uma experiência autêntica e próxima da prática instrumental

²Vídeo de demonstração acessível através da URL: <https://tinyurl.com/Hero-keys>

³GitHub: <https://github.com/JoaquimBreno/hero-keys>

real. O sistema de resposta visual inspirado em jogos como Guitar Hero gamifica o processo de aprendizado, tornando-o mais acessível e motivador para iniciantes, ao mesmo tempo que aumenta o engajamento dos usuários. Além do valor educacional, o HeroKeys apresenta potencial de desenvolvimento empreendedor significativo, pois a combinação de ferramentas de IA para transcrição musical com interface gamificada preenche uma lacuna no mercado de aprendizado musical.

O HeroKeys democratiza o ensino de piano através de transcrição automática de qualquer áudio, visualizações e gamificação fundamentada nas teorias da Autodeterminação e Flow, com feedback preciso. Limitações incluem dependência de serviços externos e sensibilidade à qualidade da gravação.

Como trabalhos futuros, pretende-se realizar estudos de caso com usuários de diferentes níveis de experiência, a fim de compreender suas interações com a aplicação, avaliar a satisfação e identificar melhorias potenciais. As evoluções planejadas incluem a implementação de interface compatível com teclado convencional do computador, permitindo a prática mesmo sem teclado musical e a ampliação da biblioteca de timbres (piano acústico, elétrico e sintetizado), enriquecendo a experiência sonora. Além disso, pretende-se fazer a incorporação da segmentação automática de músicas com o módulo Music.ai, possibilitando a identificação de seções como introdução, verso e refrão, e a prática focada em trechos específicos para facilitar o aprendizado progressivo. Tais aprimoramentos tornarão o HeroKeys uma ferramenta ainda mais completa, imersiva e motivadora para aspirantes a músicos.

REFERENCES

- [1] Marli Appel-Silva, Guilherme Welter Wendt, and Irani Iracema de Lima Argimon. 2010. A teoria da autodeterminação e as influências socioculturais sobre a identidade. *Psicol. rev. (Belo Horizonte)* 16, 2 (ago 2010), 351–369. http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-11682010000200008&lng=pt&nrm=iso Acesso em: 27 set. 2025.
- [2] Estefan Cano, Derry FitzGerald, Antoine Liutkus, Mark D. Plumbley, and Fabian Stöter. 2019. Musical source separation: An introduction. *IEEE Signal Processing Magazine* 36, 1 (2019), 31–40. <https://doi.org/10.1109/MSP.2018.2874719>
- [3] Mihaly Csikszentmihalyi. 2014. Teoria do Flow, pesquisa e aplicações. *ComCiência* 161 (set 2014). http://comciencia.scielo.br/scielo.php?script=sci_arttext&pid=S1519-76542014000700010&lng=pt&nrm=iso Acesso em: 27 set. 2025.
- [4] FastAPI. 2025. FastAPI framework, high performance, easy to learn, fast to code, ready for production. <https://fastapi.tiangolo.com/>. Acesso em: jul. 2025.
- [5] Harmonix Music Systems. 2005. Guitar Hero. Video Game. PlayStation 2.
- [6] klang.io. 2006. Transcribe Piano Music into Notes. Retrieved July 7, 2025 from <https://klang.io/piano2notes>
- [7] Qiuqiang Kong, Bochen Li, Xuchen Song, Yuan Wan, and Yuxuan Wang. 2020. High-resolution Piano Transcription with Pedals by Regressing Onsets and Offsets Times. *CoRR abs/2010.01815* (2020). arXiv:2010.01815 <https://arxiv.org/abs/2010.01815>
- [8] W. Molloy, E. Huang, and B. C. Wünsche. 2019. Mixed Reality Piano Tutor: A Gamified Piano Practice Environment. In *Proceedings of the International Conference on Electronics, Information, and Communication (ICEIC)*. Auckland, New Zealand, 1–7. <https://doi.org/10.23919/ELINFOCOM.2019.8706474>
- [9] Music.AI. 2025. AI Audio Models to Power Your Music Business. <https://music.ai/>. Acesso em: jul. 2025.
- [10] Next.js. 2025. The React Framework for the Web. <https://nextjs.org/>. Acesso em: jul. 2025.
- [11] I. G. Putranda, A. M. Yumna, Y. Rosmansyah, and Y. Sukmana. 2023. Exploring Audio Processing in Mixed Reality to Boost Motivation in Piano Learning. *IEEE Access* 11 (2023), 71194–71200. <https://doi.org/10.1109/ACCESS.2023.3293250>
- [12] Fabiola Rosendo Vignola, Nancy Dominguez González, Daniel Serna Poot, and Ernesto Vilches Lleó. 2023. Estrategia de gamificación para la enseñanza de solfeo a niños. *RIDE. Rev. Iberoam. Investig. Desarro. Educ* 14, 27 (dez 2023), e536. <https://doi.org/10.23913/ride.v14i27.1611> Epub 22-Jan-2024. Acesso em: 27 set. 2025.