



INSTITUTO FEDERAL

Sertão Pernambucano

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO SERTÃO
PERNAMBUCANO
PRÓ-REITORIA DE PESQUISA, INOVAÇÃO E PÓS-GRADUAÇÃO (PROPIP)
CAMPUS SALGUEIRO
TECNOLOGIA EM SISTEMAS PARA INTERNET**

Vinício Eufrazio Pereira da Silva

**RECONHECIMENTO DE EMOÇÕES PELA FALA: PROTÓTIPO E
MELHORIAS NO APOIO EMOCIONAL A DISTÂNCIA**

Salgueiro

2026

Vinício Eufrázio Pereira da Silva

**RECONHECIMENTO DE EMOÇÕES PELA FALA: PROTÓTIPO E
MELHORIAS NO APOIO EMOCIONAL A DISTÂNCIA**

Trabalho de Conclusão de Curso apresentado ao curso de graduação em Tecnologia em Sistemas para Internet, ofertado pelo Instituto Federal de Educação, Ciência e Tecnologia do Sertão Pernambucano, como parte dos requisitos para obtenção do título de Tecnólogo em Sistemas para Internet.

Orientador: Gustavo Freitas Sanchez

Salgueiro

2026

O trabalho de conclusão de curso “**Reconhecimento de emoções pela fala: Protótipo e melhorias no apoio emocional a distância**”, autoria de **Vinício Eufrázio Pereira da Silva**, foi submetida à Banca Examinadora, constituída pela EMEL/IFSertãoPE, como requisito parcial necessário à obtenção do título de Tecnólogo em Sistemas para Internet, outorgado pelo Instituto Federal de Educação, Ciência e Tecnologia do Sertão Pernambucano – IFSertãoPE.

Aprovado em 03 de março de 2026.

COMISSÃO EXAMINADORA

Prof. Dr. Gustavo Freitas Sanchez – IFSertãoPE
(Presidente)

Prof. Dr. Marcelo Anderson Batista dos Santos – IFSertãoPE
(1º Examinador)

Prof. Allanyo Antonio Silva Santos – IFSertãoPE
(2º Examinador)

Dados Internacionais de Catalogação na Publicação (CIP)

P436 Pereira da Silva, Vinicio Eufrazio.

Reconhecimento de emoções pela fala: Protótipo e melhorias no apoio emocional a distância / Vinicio Eufrazio Pereira da Silva. - Salgueiro, 2026.
33 f. : il.

Trabalho de Conclusão de Curso (Sistemas para Internet) -Instituto Federal de Educação, Ciência e Tecnologia do Sertão Pernambucano, Campus Salgueiro, 2026.
Orientação: Prof. Dr. Gustavo Freitas Sanchez.

1. Inteligência artificial. 2. Saúde mental. 3. Inteligência artificial. 4. Reconhecimento das emoções. I. Título.

CDD 006.3

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, por ter sido meu alicerce e meu refúgio em todos os momentos desta caminhada. Foi por meio da fé, da força e do apoio emocional que encontrei perseverança para superar os desafios enfrentados ao longo do curso e, especialmente, durante a elaboração deste trabalho.

Expresso minha profunda gratidão à equipe da COINFO, que esteve presente em minha trajetória ao longo dos sete anos de formação no IF, desde o ensino médio até a graduação. Sinto-me verdadeiramente próximo de cada professor, de cada ensinamento compartilhado e de cada experiência vivenciada nesse período, que contribuíram significativamente para minha formação acadêmica e pessoal.

Agradeço, de maneira especial, ao meu orientador, Professor Gustavo, pelo apoio constante, pelas valiosas contribuições, pelas ideias compartilhadas e, sobretudo, pela paciência e dedicação durante todo o processo de desenvolvimento deste trabalho.

Sua orientação foi essencial para a concretização deste projeto. Estendo meus agradecimentos à minha família, pelo incentivo, compreensão e apoio incondicional em todos os momentos. Foram fundamentais para que eu pudesse chegar até aqui.

Por fim, agradeço a mim mesmo, pela persistência, disciplina e determinação em concluir esta etapa tão aguardada e significativa da minha vida. Este trabalho representa não apenas uma conquista acadêmica, mas também a realização de um ciclo de crescimento e superação.

RESUMO

Este trabalho aborda o desenvolvimento de uma plataforma digital de apoio emocional integrada a um sistema de reconhecimento automatizado de emoções por meio da fala, tendo como objetivo propor uma solução complementar ao atendimento psicológico tradicional diante do aumento dos casos de ansiedade e depressão e das limitações estruturais do acesso à saúde mental no Brasil. A pesquisa adotou abordagem qualitativa, exploratória e bibliográfica, fundamentada na análise de documentos institucionais, artigos científicos e estudos acadêmicos sobre saúde mental, telepsicologia, inteligência artificial emocional e reconhecimento de emoções por voz. Além da revisão teórica e da análise de plataformas digitais existentes, foram realizados testes experimentais em ambiente controlado com modelo baseado em rede neural densa (DNN), utilizando base de dados rotulada para classificação emocional, com avaliação por matriz de confusão. Como resultado, foi desenvolvido o conceito estrutural de uma plataforma que integra múltiplas camadas de suporte: diário emocional digital com opção de desabafo anônimo ou identificado, jogos interativos voltados à regulação emocional, monitoramento longitudinal de estados emocionais e sistema de reconhecimento emocional por voz com classificação probabilística. Os testes indicaram viabilidade técnica preliminar do modelo de análise da fala, embora com limitações relacionadas à base de dados e à validação em larga escala. Conclui-se que a originalidade do trabalho reside na integração sistêmica entre monitoramento automatizado, expressão emocional, recursos lúdicos e acompanhamento contínuo, configurando um ecossistema digital de suporte preventivo e complementar ao cuidado profissional, com atenção a princípios éticos como privacidade, consentimento informado e não substituição do atendimento psicológico especializado.

Palavras-Chave: Saúde mental; Inteligência Artificial; Reconhecimento de emoções pela fala.

ABSTRACT

This study presents the conceptual and structural development of a digital emotional support platform integrated with an automated speech-based emotion recognition system. The objective is to propose a complementary technological solution to traditional psychological care, considering the increasing prevalence of anxiety and depression and the structural limitations of mental health services in Brazil. The research adopted a qualitative, exploratory, and bibliographic approach, based on the analysis of institutional reports, scientific articles, and academic studies related to mental health, telepsychology, emotional artificial intelligence, and speech emotion recognition. In addition to the theoretical review and analysis of existing digital platforms, controlled experimental tests were conducted using a dense neural network (DNN) model trained on labeled emotional speech datasets, with performance evaluated through a confusion matrix. The results led to the conceptual development of an integrated platform combining multiple support layers: a digital emotional diary with anonymous or identified expression, interactive games aimed at emotional regulation, longitudinal emotional monitoring, and probabilistic speech-based emotion classification. Preliminary findings indicate technical feasibility of the speech recognition component, although limitations remain regarding dataset representativeness and large-scale validation. The originality of this work lies in the systemic integration of automated monitoring, emotional expression tools, gamified coping strategies, and continuous tracking within a preventive digital ecosystem designed to complement — rather than replace — professional psychological care, while respecting ethical principles such as data privacy, informed consent, and responsible AI use.

Keywords: Mental health; Artificial Intelligence; Speech emotion recognition.

LISTA DE FIGURAS

Figura 01 - Tela inicial Zenklub

Figura 02 - Tela inicial Vittude

Figura 03 - Tela inicial Wysa

Figura 04 - Dashboard plataforma desenvolvida aplicando as melhorias sugeridas do trabalho

Figura 05 - Matriz de confusão do modelo de identificação de voz usado

Figura 06 - Informações do treinamento (precision, recall, f1-score)

Figura 07 - Exemplo de como é a análise de emoção através do protótipo de plataforma desenvolvido

LISTA DE ABREVIATURAS E SIGLAS

IA - Inteligência Artificial

CNN / CNNs - Rede Neural Convolucional (Convolutional Neural Network)

LSTM - Memória de Longo e Curto Prazo (Long Short-Term Memory)

OPAS - Organização Pan-Americana da Saúde

RAVDESS - Ryerson Audio-Visual Database of Emotional Speech and Song

UFAM - Universidade Federal do Amazonas

UFPB - Universidade Federal da Paraíba

UNESP - Universidade Estadual Paulista

SUS - Sistema Único de Saúde

OMS - Organização Mundial da Saúde

SUMÁRIO

1 INTRODUÇÃO	11
2 REFERENCIAL TEÓRICO	12
3 METODOLOGIA	15
4 ANÁLISE DOS DADOS (RESULTADOS E DISCUSSÕES)	27
5 CONCLUSÕES (CONSIDERAÇÕES FINAIS)	31
REFERÊNCIAS	32

1 INTRODUÇÃO

Nos últimos anos é possível observar o crescimento de tópicos relacionados a transtornos emocionais, aumento nos diagnósticos de ansiedade e depressão em diversas faixas etárias, de jovens adolescentes até adultos. Segundo o Ministério da Saúde (2022), a prevalência de depressão na rede de atenção primária no Brasil é alta, e há uma grande lacuna no atendimento especializado para saúde mental. Além disso, a pandemia de COVID-19 acabou contribuindo para um aumento significativo nos diagnósticos de prevalência nessas condições, cerca de 25%, como aponta a OPAS - Organização Pan-Americana da Saúde (2022).

No Brasil a telepsicologia é um método onde o paciente e psicólogo não se encontram pessoalmente, mas sim por meio de videochamadas, e-mails, ligações, mensagens de texto, entre outros, tem sido bastante usada como alternativa para atender a demanda por cuidados emocionais de acordo com a Saúde Digital Brasil (2023).

Paralelamente aos problemas apresentados anteriormente, o avanço da inteligência artificial (IA) abre novas possibilidades para a saúde mental também, isto é, a capacidade de algoritmos interpretarem estados emocionais. Para o Brasil, estudos recentes investigam a aplicação da IA para suporte psicológico: por exemplo, chatbots terapêuticos estão sendo usados como aliados de apoio 24h para pacientes em um projeto comandado pelo professor José Humberto da Silva Filho da Faculdade de Psicologia, da Universidade Federal do Amazonas (2024).

Ainda a fim de complementar nosso trabalho, existem pesquisas acadêmicas que mostram como a voz pode ser usada para detectar emoções através da IA. Uma pesquisa realizada na UNESP em 2022 tratou de redes neurais convolucionais (CNNs) para classificar estados emocionais a partir da voz.

Dado esse cenário apresentado anteriormente, o presente trabalho propõe o desenvolvimento de uma plataforma online de apoio emocional integrada a um sistema de reconhecimento automático de emoções por voz que faz uso de uma IA treinada para identificar padrões emocionais a partir de gravações dos usuários.

Essa solução tem o objetivo não de substituir a psicoterapia tradicional, mas oferecer um ambiente complementar, onde é possível monitorar o dia a dia desse paciente, dar orientações iniciais e estimular o autocuidado.

2 REFERENCIAL TEÓRICO

2.1 Saúde mental e Transtornos emocionais

A saúde mental é um estado de bem-estar em que o indivíduo é capaz de lidar com as dificuldades na vida, ser produtivo e contribuir para com a comunidade. Porém, transtornos emocionais como depressão e ansiedade são cada vez mais frequentes, o que vai contra o bem-estar mental. De acordo com o ministério da Saúde (2022), a depressão apresenta alta prevalência na atenção primária à saúde no Brasil ou seja, um grande aumento de diagnósticos em postinhos de bairro e unidades básicas de saúde de doenças relacionadas ao emocional, mostrando que há a necessidade de estratégias que ampliem o acesso ao cuidado psicológico.

A pandemia de COVID-19 sem dúvidas agrava esse cenário, intensificando sentimentos de medo, insegurança através do isolamento social e pelo fato de ser uma situação que ninguém havia vivenciado. De acordo com a Organização Pan-Americana de Saúde (2022), houve um aumento de aproximadamente 25% nos casos de depressão e ansiedade durante a pandemia. Isso é, fica evidente a necessidade de soluções que possam atuar de forma preventiva e principalmente, complementar ao atendimento tradicional da saúde mental.

2.2 Telepsicologia e Saúde Digital

A telepsicologia é um termo que consiste na prestação de serviços psicológicos através de tecnologias digitais, vídeo chamadas, mensagens, e-mails e plataformas online. Esse modelo é útil principalmente por superar barreiras geográficas. Atrelado a isso também a telepsiquiatria, que também é um termo que consiste na prestação de serviços psiquiátricos através de tecnologias digitais está sendo bem aderido.

No Brasil, a adoção da saúde digital cresceu significativamente nos últimos anos. Segundo a Saúde Digital Brasil (2024), houve um aumento expressivo no número de atendimentos de telepsicologia e telepsiquiatria, esses dados indicam que plataformas online têm potencial para complementar os serviços presenciais oferecendo suporte a distância além de ampliar o alcance do atendimento psicológico e psiquiátrico.

2.3 Inteligência Artificial Aplicada à Saúde Mental

A IA tem sido amplamente estudada como ferramenta de apoio a saúde mental, principalmente em análise de padrões comportamentais e emocionais. Com capacidade computacional suficiente para reconhecer, interpretar e responder a estados emocionais humanos, com dados como texto, imagem e voz.

Segundo reportagem do jornal O Globo (2025), a chamada "IA emocional" levanta discussões sobre ética, privacidade e confiabilidade, porém também apresenta grande potencial em como os usuários interagem com tecnologias de suporte emocional. No Brasil, universidades têm explorado essa tecnologia como recurso complementar ao cuidado psicológico e não como "única e melhor forma de cuidar do psicológico".

Um exemplo de referência é o projeto desenvolvido na Universidade Federal do Amazonas (UFAM), que utiliza chatbots terapêuticos baseados em IA para oferecer suporte emocional contínuo aos usuários. Conforme a UFAM (2024), esses sistemas atuam como aliados no atendimento psicológico, fornecendo escuta ativa, orientações iniciais e encaminhamentos quando necessário, sem substituir o acompanhamento profissional.

2.4 Reconhecimento das emoções através da voz

A voz humana carrega informações aproximadas sobre o estado emocional de uma pessoa, o que torna isso possível é a análise de alguns atributos como o tom de voz principalmente, além do ritmo, intensidade e pausas, por exemplo. Com base nisso, modelos de aprendizado de máquinas têm sido desenvolvidos para classificar possíveis estados emocionais.

Pesquisas acadêmicas realizadas na UNESP, por Peixoto e Linhares (2023), usou redes neurais convolucionais (CNNs) para o reconhecimento de emoções na voz, obtendo bons resultados mesmo em ambientes com ruído, e em português! já que muitos modelos pré treinados em sua maioria são encontrados em inglês. Esse trabalho evidencia que técnicas baseadas em deep learning são capazes de extrair padrões emocionais através da voz humana.

Semelhante à pesquisa anterior, Aguiar(2024) em um Trabalho de Conclusão de Curso, na UFPB, empregou as arquiteturas CNN e LSTM para o reconhecimento de emoções em áudios em português, usando a base de dados emoUERJ. O que reforça a eficácia da combinação entre redes CNNs e modelos para análise da voz.

2.5 Plataformas digitais de apoio emocional

Plataformas digitais que fazem uso de recursos como tecnologias de IA para auxiliar no cuidado psicológico permitem o monitoramento contínuo do estado emocional dos usuários, podendo oferecer feedbacks, orientações e estímulo ao autocuidado.

Diferentemente da psicoterapia tradicional, essas soluções NÃO têm como objetivo substituir o profissional de saúde mental, mas atuar como ferramentas complementares. Com integração com reconhecimento de emoções através da voz, essas plataformas podem identificar padrões emocionais recorrentes e auxiliar o usuário, direcionando-o a um profissional da área para compreender melhor seu próprio estado psicológico.

Nesse contexto, este trabalho fundamenta-se nos estudos apresentados para discutir e analisar o uso de plataformas digitais de apoio psicológico que utilizam a análise automatizada da fala, alinhando-se às tendências atuais da saúde digital e da aplicação da inteligência artificial ao bem-estar emocional.

3 METODOLOGIA

Este trabalho adota uma abordagem metodológica de caráter qualitativo, exploratório e bibliográfico, com o objetivo de analisar o cenário atual da saúde mental no contexto digital e identificar possibilidades de aplicação da inteligência artificial como ferramenta de apoio emocional. A metodologia foi estruturada em etapas que permitem compreender o problema, analisar soluções existentes e propor diretrizes conceituais para iniciativas tecnológicas voltadas ao suporte psicológico.

3.1 Tipo de pesquisa

A pesquisa caracteriza-se como exploratória, pois busca ampliar o entendimento sobre o uso de tecnologias digitais e da inteligência artificial no apoio à saúde mental, um campo ainda em expansão e com constantes transformações. Segundo esse tipo de abordagem, o foco não está na validação estatística, mas na compreensão do fenômeno e na identificação de oportunidades de melhoria e inovação.

Além disso, trata-se de uma pesquisa bibliográfica, fundamentada na análise de artigos científicos, documentos institucionais, relatórios de organizações de saúde e trabalhos acadêmicos que abordam saúde mental, saúde digital, telepsicologia e inteligência artificial emocional.

3.2 Levantamento e análise do cenário atual da Saúde Mental

Nesta etapa, realizou-se uma análise dos dados oficiais, relatórios institucionais e documentos publicados por órgãos como o Ministério da Saúde e A OPAS, que são os que apresentam informações com propriedade acerca de transtornos como ansiedade e depressão, bem como as limitações estruturais existentes, na realização de atendimento especializado em saúde mental. Ressalta-se ainda que a proposta deste trabalho é datar, analisar e propor possíveis melhorias para este tipo de plataforma.

Foram levados em consideração também os estudos realizados em no período pandêmico da COVID-19. A análise desses cuidados permitiu

identificar lacunas importantes no acesso ao cuidado psicológico, essa análise associa-se a o avanço tecnológico neste período, pois foi o que possibilitou o surgimento/melhorias de ideias, tecnologias e plataformas de auxílio mental, parte dessas lacunas são a insuficiência de profissionais para atender à demanda crescente, barreiras geográficas e a dificuldade de acompanhamento contínuo dos pacientes. Esses fatores evidenciam a necessidade de soluções complementares que possam auxiliar no suporte emocional, sem substituir o atendimento profissional tradicional.

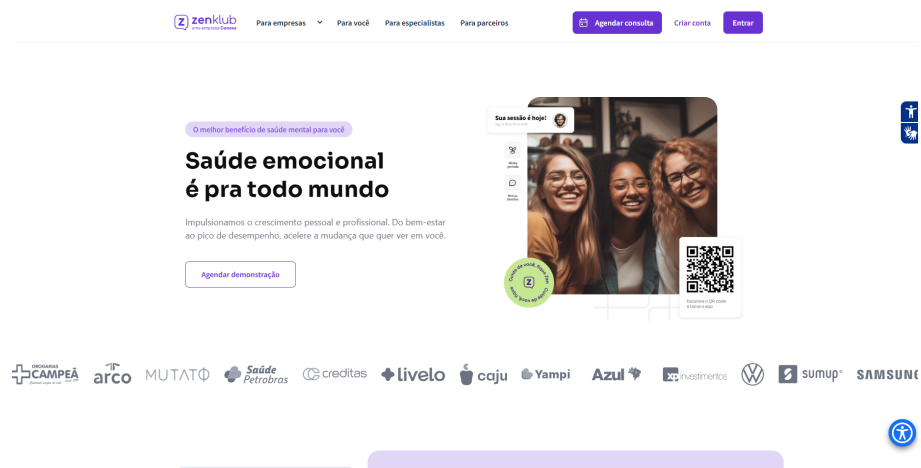
3.3 Análise de soluções digitais e iniciativas existentes

Após a compreensão do cenário da saúde mental, foi realizada uma análise de soluções digitais já existentes voltadas ao apoio emocional e psicológico. Essa etapa teve como objetivo identificar como a tecnologia vem sendo utilizada para suprir, ainda que parcialmente, as limitações do atendimento presencial.

Foram analisadas plataformas de destaque no cenário brasileiro, com ênfase em soluções consolidadas de atendimento psicológico online, como o Zenklub e a Vittude. A escolha dessas plataformas se deu por sua relevância no mercado nacional e por apresentarem modelos distintos de atuação no suporte emocional digital.

- Zenklub é uma plataforma brasileira de saúde emocional que oferece atendimento psicológico e psiquiátrico online, conectando usuários a psicólogos, psicanalistas, terapeutas e psiquiatras por meio de vídeo chamadas, chat e troca de mensagens.

Figura 01 - Tela inicial Zenklub



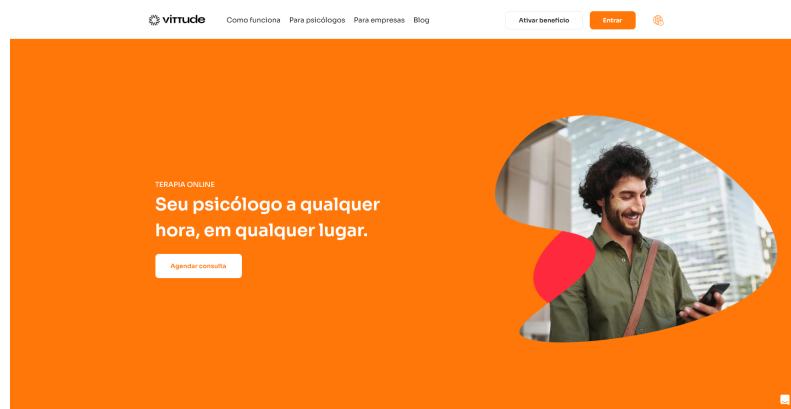
Fonte: <https://zenklub.com.br>

O Zenklub apresenta um conjunto de funcionalidades voltadas a telepsicologia e telepsiquiatria, incluindo agendamento de sessões, realização de videochamadas, troca de mensagens assíncronas entre usuário e profissional e disponibilização de conteúdos educativos relacionados ao bem-estar emocional. Essas funcionalidades contribuem para a ampliação do acesso ao cuidado psicológico, especialmente em contextos nos quais o atendimento presencial é limitado por fatores geográficos ou estruturais.

No entanto observa-se que a atuação da plataforma ocorre principalmente de forma reativa, ou seja, concentrando o acompanhamento emocional nos momentos de interação direta entre usuário e profissional. Como oportunidade de melhoria, destaca-se a possibilidade de implementação de mecanismos de monitoramento contínuo do estado emocional do usuário, por meio de check-ins periódicos, questionários dinâmicos ou registros de humor ao longo do tempo. Esse tipo de acompanhamento permitiria uma visão mais ampla e longitudinal do estado emocional do indivíduo.

Outra possibilidade relevante de aprimoramento seria a integração de recursos de inteligência artificial capazes de identificar padrões emocionais a partir das interações do usuário com a plataforma, como frequência de uso, variações nos relatos emocionais e histórico de atendimentos. A análise desses dados poderia auxiliar tanto os profissionais quanto os próprios usuários na identificação precoce de sinais de agravamento emocional, contribuindo para intervenções mais rápidas e assertivas.

Figura 02 - Tela inicial Vittude



Fonte: <https://vittude.com>

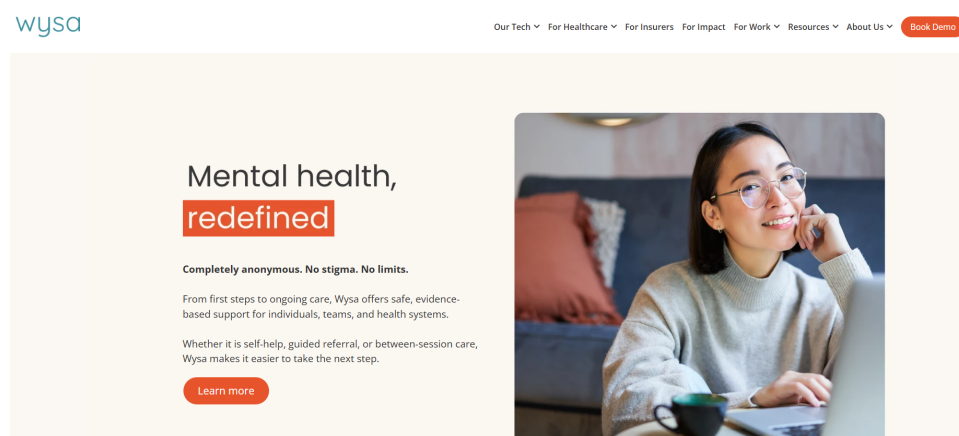
De forma semelhante, a plataforma Vittude se destaca pela utilização de questionários iniciais como estratégia de triagem, direcionando o usuário ao profissional mais adequado de acordo com suas demandas emocionais. Essa abordagem representa um avanço em relação a modelos tradicionais de encaminhamento e é particularmente relevante em contextos corporativos, nos quais a escala de atendimento é elevada.

No entanto, após a etapa inicial de triagem, o acompanhamento emocional tende a se tornar mais estático, dependendo essencialmente das sessões terapêuticas realizadas. Como oportunidade de melhoria, identifica-se a adoção de ferramentas de acompanhamento longitudinal que permitam o registro contínuo do estado emocional do usuário ao longo do tempo, inclusive fora dos momentos de consulta.

Outra limitação observada refere-se à ausência de recursos interativos automatizados, como chatbots terapêuticos de apoio emocional imediato. A implementação desses recursos poderia oferecer suporte inicial em momentos de vulnerabilidade emocional, funcionando como uma camada complementar de acolhimento, sem substituir o atendimento humano especializado.

O Wysa é uma plataforma internacional de saúde mental com abordagem de ecossistema integrado, que combina chatbot baseado em inteligência artificial, técnicas de terapia cognitivo-comportamental (TCC), exercícios de mindfulness, monitoramento de humor e encaminhamento opcional a profissionais humanos. Diferentemente das plataformas de teleconsulta, o Wysa atua de forma proativa e contínua, oferecendo suporte emocional imediato independentemente da disponibilidade de um terapeuta.

Figura 03: Wysa tela inicial



Fonte: <https://www.wysa.com>

Entre seus principais recursos, destacam-se: check-ins diários de estado emocional, histórico longitudinal de variações de humor, sessões guiadas de técnicas cognitivo-comportamentais e integração entre suporte automatizado e atendimento profissional quando necessário. A plataforma recebeu em 2025 o status de Breakthrough Device da FDA norte-americana, reconhecimento que atesta seu potencial de impacto em saúde mental (Inkster et al., 2018).

Como limitação central, o Wysa não incorpora reconhecimento emocional por meio da análise acústica da fala, sendo o texto escrito o único canal de coleta de dados emocionais. Adicionalmente, a plataforma não foi desenvolvida para o contexto brasileiro, apresentando barreiras linguísticas e culturais para sua adoção no país.

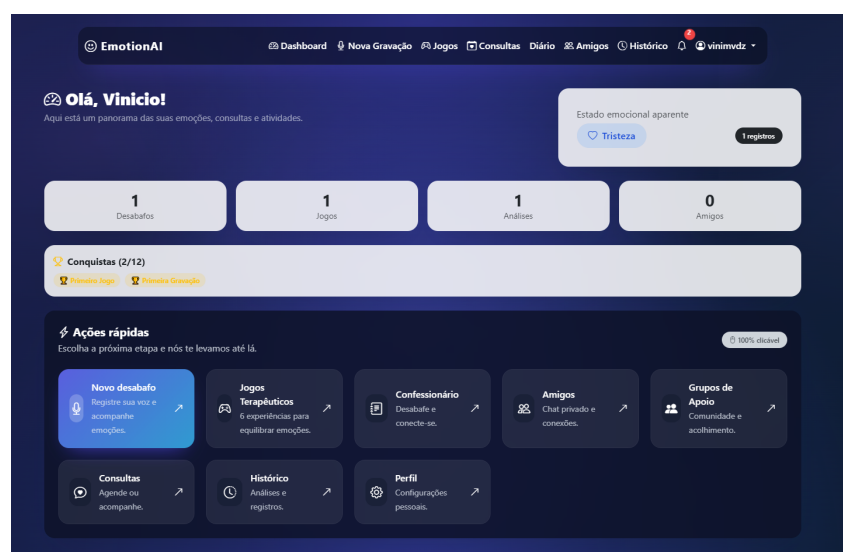
3.4 Identificação de Possíveis Melhorias e Oportunidades no Setor

A análise das três plataformas evidencia dois modelos distintos de suporte emocional digital. Zenklub e Vittude representam o modelo de teleconsulta, centrado na mediação tecnológica da relação entre paciente e profissional, com atuação predominantemente reativa. Wysa representa um modelo de ecossistema digital, caracterizado pela integração de múltiplos recursos de suporte, monitoramento contínuo e uso estrutural de inteligência artificial.

A plataforma proposta neste trabalho alinha-se ao segundo modelo, incorporando elementos como diário emocional digital, monitoramento longitudinal e reconhecimento emocional por voz. Em relação às plataformas de ecossistema

analisadas, a principal contribuição diferencial reside em dois aspectos: a integração da análise acústica da fala como mecanismo complementar de identificação do estado emocional — funcionalidade ausente em todas as plataformas analisadas — e o desenvolvimento orientado ao contexto brasileiro, com consideração às especificidades linguísticas e culturais da população nacional.

Figura 04 - Dashboard plataforma desenvolvida aplicando as melhorias sugeridas do trabalho



Fonte: (próprio autor)

3.5 Pipeline do modelo de reconhecimento emocional por voz

Para o treinamento e avaliação do modelo de reconhecimento emocional, foi utilizada a base de dados RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song), desenvolvida pela Ryerson University e disponibilizada publicamente na plataforma Kaggle. O RAVDESS é amplamente adotado na literatura de reconhecimento emocional por voz por apresentar gravações de alta qualidade com rotulação emocional padronizada e processo de coleta com atores profissionais treinados como datam Livingstone, Russo (2018).

A base contém gravações de 24 atores — 12 homens e 12 mulheres — expressando oito categorias emocionais: neutro, calmo, feliz, triste, raiva, medo, nojo e surpresa. Cada emoção foi registrada em dois níveis de intensidade, normal e forte, resultando em 1.440 arquivos de fala. Para este trabalho, foram utilizados

exclusivamente os arquivos de modalidade speech, excluindo-se os arquivos de canto.

É importante reconhecer que o RAVDESS foi produzido em língua inglesa com atores de origem norte-americana, o que representa uma limitação relevante para aplicações no contexto brasileiro. A expressão emocional é influenciada por aspectos culturais, linguísticos e prosódicos específicos de cada idioma; modelos treinados exclusivamente nessa base podem, portanto, apresentar menor precisão quando aplicados a falantes nativos de português. Essa limitação é assumida no presente trabalho e aponta para a necessidade de desenvolvimento de bases equivalentes em português brasileiro para avaliações futuras — lacuna já identificada por Peixoto e Linhares (2023) e Aguiar (2024).

Como referência técnica para esse tipo de abordagem, foi considerado o uso de bases públicas de dados amplamente utilizadas na literatura de reconhecimento emocional por voz, como o RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song), disponibilizado publicamente na plataforma Kaggle. O RAVDESS contém gravações de fala e canto com diferentes categorias emocionais previamente rotuladas, sendo frequentemente empregado em pesquisas para treinamento e validação de modelos de classificação emocional baseados em aprendizado de máquina.

3.5.1 Arquitetura do modelo e protocolo de treinamento

O modelo de classificação emocional foi implementado utilizando uma Rede Neural Densa (DNN — Dense Neural Network), também denominada Multilayer Perceptron (MLP), por meio da API Sequential do framework TensorFlow/Keras. A escolha por uma arquitetura densa, em oposição a uma rede convolucional (CNN), justifica-se pelo formato dos dados de entrada: vetores unidimensionais de 40 características MFCC agregadas, sem preservação de estrutura espacial ou temporal que exigiria operações de convolução.

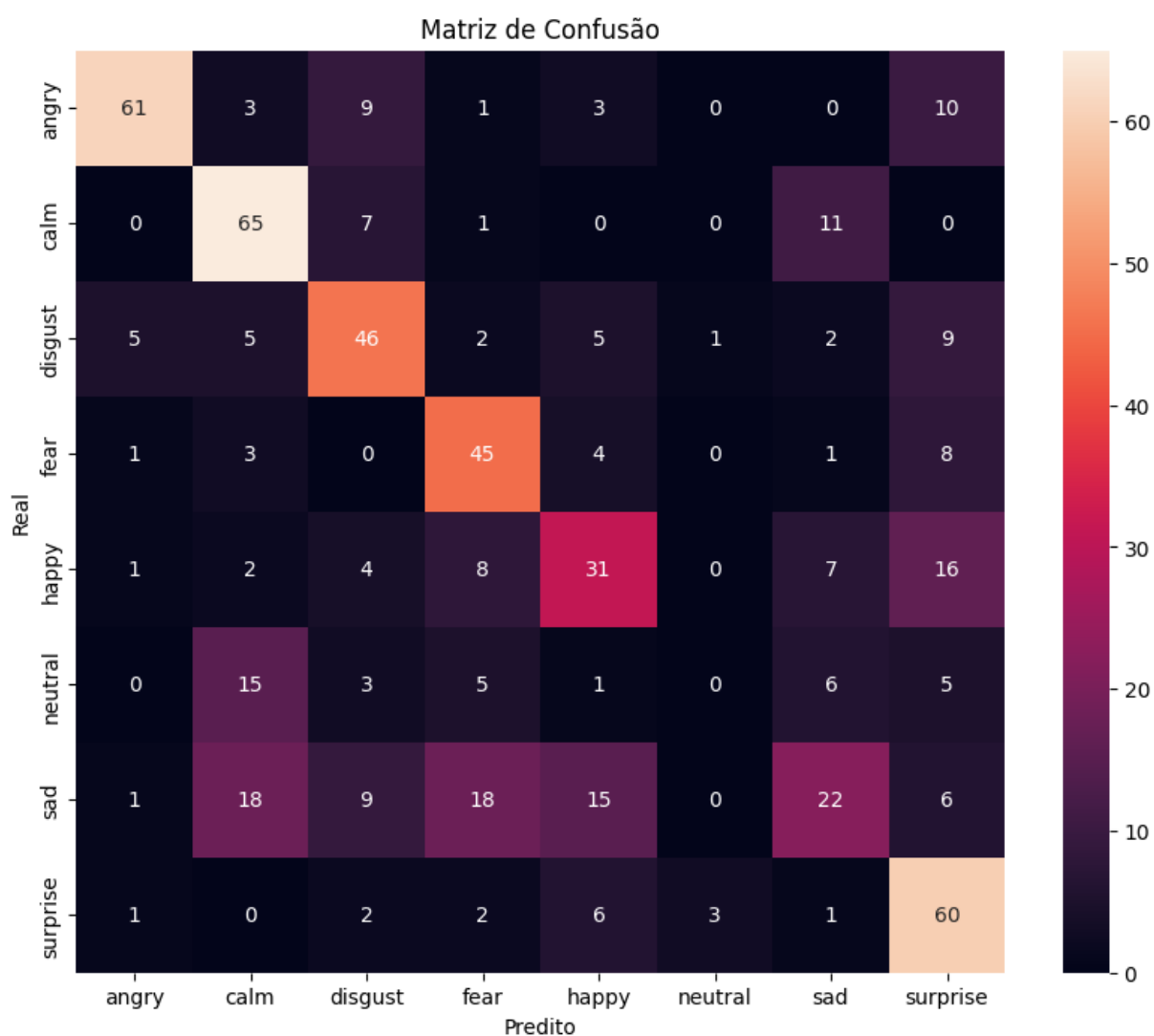
A arquitetura é composta por quatro camadas: uma camada de entrada com 256 neurônios e ativação ReLU, seguida de Dropout com taxa de 0,3 para regularização; uma segunda camada oculta com 128 neurônios e ativação ReLU, também com Dropout de 0,3; uma terceira camada com 64 neurônios e ativação ReLU; e uma camada de saída com 8 neurônios — correspondentes às oito classes emocionais — com função de ativação softmax, que retorna uma distribuição de probabilidade sobre as categorias. O modelo foi compilado com o otimizador Adam (taxa de aprendizado de 0,001) e função de perda de entropia cruzada categórica (`categorical_crossentropy`).

O conjunto de dados foi dividido em 80% para treinamento e 20% para teste, com `random_state=42` para reprodutibilidade. O treinamento foi conduzido por 50 épocas com `batch_size` de 16 e monitoramento simultâneo da perda e acurácia no conjunto de validação (`X_test`, `y_test`).

3.5.2 Avaliação do modelo e interpretação dos resultados

O desempenho do modelo foi avaliado por meio do relatório de classificação e da matriz de confusão, gerados com as funções `classification_report` e `confusion_matrix` da biblioteca `scikit-learn`. O conjunto de teste totalizou 576 amostras distribuídas entre as oito classes emocionais.

Figura 05 - Matriz de confusão do modelo de identificação de voz usado



Fonte: (próprio autor)

O modelo obteve acurácia geral de 57%, com F1-score médio ponderado de 0,55. A análise por classe revela desempenhos heterogêneos: angry apresentou o melhor resultado, com F1-score de 0,78 e precisão de 0,87, o que se explica pelas características acústicas marcantes dessa emoção, como elevada intensidade vocal, frequência fundamental aumentada e ritmo acelerado. A categoria surprise também obteve bom recall (0,80), com 60 dos 75 exemplos corretamente identificados.

Figura 06: Informações do treinamento (precision, recall, f1-score)

	precision	recall	f1-score	support
angry	0.87	0.70	0.78	87
calm	0.59	0.77	0.67	84
disgust	0.57	0.61	0.59	75
fear	0.55	0.73	0.62	62
happy	0.48	0.45	0.46	69
neutral	0.00	0.00	0.00	35
sad	0.44	0.25	0.32	89
surprise	0.53	0.80	0.63	75
accuracy			0.57	576
macro avg	0.50	0.54	0.51	576
weighted avg	0.54	0.57	0.55	576

Fonte: (próprio autor)

Por outro lado, a categoria neutral apresentou desempenho nulo — precisão, recall e F1-score iguais a zero —, indicando que o modelo não conseguiu aprender padrões discriminativos para essa classe. Esse resultado é compreensível, pois o estado neutro caracteriza-se justamente pela ausência de marcadores emocionais salientes, tornando-o acusticamente semelhante a várias outras categorias. A classe sad também apresentou desempenho baixo ($F1 = 0,32$), com alta dispersão nas predições, sendo frequentemente confundida com calm, fear e happy, conforme evidenciado na matriz de confusão (Figura 04).

Esses resultados são consistentes com a literatura de reconhecimento emocional por voz, que reporta sistematicamente maior dificuldade na classificação de emoções com baixa intensidade vocal, como neutro e triste (Peixoto; Linhares, 2023). Ressalta-se que os testes foram conduzidos em ambiente controlado, com dados da mesma origem e idioma do treinamento; o desempenho tende a ser inferior em condições reais com falantes de português brasileiro, reforçando a necessidade de validação com bases culturalmente representativas em trabalhos futuros.

A viabilidade técnica da abordagem foi demonstrada para as classes emocionais de maior intensidade expressiva, enquanto as limitações identificadas fornecem diretrizes concretas para o aprimoramento do modelo: ampliação e diversificação da base de dados, inclusão de bases em português, técnicas de balanceamento de

classes (como oversampling da classe neutral) e experimentação com arquiteturas mais complexas, como CNNs aplicadas a espectrogramas mel ou modelos de atenção temporal.

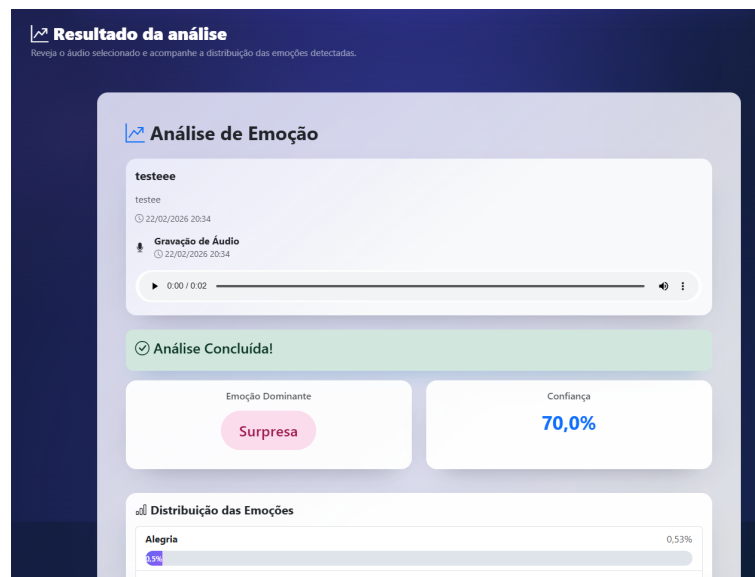
Ressalta-se que o uso desse tipo de base tem caráter experimental e demonstrativo, servindo como exemplo metodológico para pesquisadores e desenvolvedores que pretendam aplicar técnicas semelhantes. O objetivo não é afirmar que modelos treinados com o RAVDESS representam fielmente a diversidade emocional da população brasileira, mas sim demonstrar a viabilidade técnica da abordagem de reconhecimento emocional por meio da fala.

Com o objetivo de verificar a viabilidade prática do reconhecimento emocional por meio da fala, foram realizados testes experimentais internos utilizando modelo treinado com base em arquitetura de rede neural densa (DNN), aplicados sobre base de dados rotulada para classificação emocional.

Ressalta-se que os testes foram conduzidos em ambiente controlado, para fins acadêmicos e demonstrativos, não representando ainda um experimento validado em larga escala ou aplicado a amostras populacionais brasileiras diversificadas. Dessa forma, os resultados apresentados possuem caráter preliminar e podem variar conforme ampliação da base de dados, ajustes de hiperparâmetros e adaptação cultural do modelo.

A avaliação do desempenho foi realizada por meio de métricas clássicas de aprendizado de máquina, com ênfase na matriz de confusão, que permite visualizar o desempenho do classificador em cada categoria emocional.

Figura 07 - Exemplo de como é a análise de emoção através do protótipo de plataforma desenvolvido



Fonte: (próprio autor)

A utilização de datasets estruturados como o RAVDESS permite que pesquisadores realizem experimentações controladas, validem arquiteturas de redes neurais e comparem resultados com trabalhos anteriores da literatura. Contudo, também evidencia uma oportunidade de melhoria importante: a necessidade de desenvolvimento e ampliação de bases de dados em língua portuguesa e em contextos culturais brasileiros, visando maior precisão e adequação dos modelos à realidade nacional.

Além dos aspectos técnicos relacionados ao reconhecimento emocional, também foram considerados elementos não técnicos fundamentais, como a importância de garantir que soluções digitais atuem de forma ética, transparente e complementar ao acompanhamento psicológico profissional. A implementação de modelos de inteligência artificial deve observar princípios como privacidade de dados, consentimento informado e limitação de escopo, evitando interpretações determinísticas ou diagnósticos automatizados.

Dessa forma, as melhorias identificadas não se restringem apenas ao aprimoramento tecnológico, mas também envolvem a integração responsável dessas ferramentas ao cuidado em saúde mental, priorizando o suporte humanizado

e o uso da inteligência artificial como instrumento auxiliar — e não substitutivo — da atuação profissional.

3.6 Uso da Inteligência Artificial neste trabalho

A inteligência artificial também foi utilizada neste trabalho como ferramenta de apoio ao processo de pesquisa e levantamento teórico. Seu uso concentrou-se na organização, análise e sistematização de informações provenientes de artigos científicos, documentos institucionais e trabalhos acadêmicos relacionados ao tema.

Por meio do auxílio da IA, foi possível identificar referências relevantes, agrupar estudos por temática, comparar abordagens metodológicas e reconhecer tendências recentes no uso de inteligência artificial aplicada à saúde mental e ao reconhecimento de emoções por meio da fala. Esse recurso contribuiu para otimizar o processo de pesquisa bibliográfica, ampliando o alcance das fontes analisadas e auxiliando na construção de uma base teórica consistente.

Ressalta-se que a inteligência artificial foi empregada exclusivamente como ferramenta de apoio à pesquisa, não substituindo a análise crítica do autor, que foi responsável pela seleção, interpretação e validação das informações utilizadas neste trabalho.

4 ANÁLISE DOS DADOS (RESULTADOS E DISCUSSÕES)

4.1 Resultados do desenvolvimento da plataforma

O presente trabalho resultou no desenvolvimento conceitual e estrutural de uma plataforma digital de apoio emocional integrada a múltiplos recursos tecnológicos voltados ao acompanhamento contínuo do estado emocional do usuário.

Diferentemente de soluções centradas exclusivamente em teleatendimento, a plataforma proposta incorpora diferentes camadas de suporte, tais como: sistema de reconhecimento emocional por meio da fala, diário emocional digital, opção de desabafo anônimo ou identificado, jogos interativos com foco em estímulo cognitivo e regulação emocional, além de registro longitudinal dos estados emocionais.

Essa abordagem integrada possibilita que o usuário interaja com o sistema de maneira ativa, contínua e personalizada, ampliando o potencial preventivo da solução proposta.

4.2 Diário emocional e funcionalidade de desabafo

Um dos principais resultados do desenvolvimento da plataforma foi a implementação do diário emocional digital, no qual o usuário pode registrar sentimentos, pensamentos e experiências cotidianas.

A funcionalidade permite que o desabafo seja realizado tanto de forma identificada quanto de forma totalmente anônima. A possibilidade de anonimato representa um diferencial relevante, pois reduz barreiras relacionadas ao julgamento social, à vergonha ou ao receio de exposição, fatores frequentemente associados à resistência em buscar ajuda psicológica.

Além disso, os registros realizados no diário podem ser utilizados para monitoramento da frequência emocional, identificação de padrões recorrentes e apoio ao acompanhamento profissional, quando autorizado pelo próprio usuário. Esse recurso amplia o engajamento e promove o autoconhecimento, funcionando como ferramenta complementar de suporte emocional.

4.3 Jogos interativos como ferramenta de regulação emocional

A plataforma desenvolvida incorpora, em caráter auxiliar, recursos interativos com finalidade de estímulo cognitivo e distração emocional. A inclusão desse componente foi motivada por estudos que apontam o potencial de atividades lúdicas digitais como estratégia complementar de regulação emocional, especialmente em momentos de estresse agudo (Fleming et al., 2017).

No contexto da plataforma, os recursos interativos foram concebidos como mecanismos de interrupção de ciclos de pensamento negativo e promoção de engajamento do usuário, funcionando como uma camada de suporte leve entre os módulos de registro emocional e monitoramento por voz.

Reconhece-se, no entanto, que a relação entre mecânicas de jogo específicas e benefícios terapêuticos mensuráveis é complexa e demanda validação empírica que extrapola o escopo metodológico deste trabalho. A fundamentação clínica desse componente não foi investigada com profundidade equivalente à dos demais módulos da plataforma, o que constitui uma limitação a ser endereçada em trabalhos futuros.

Recomenda-se que pesquisas subsequentes investiguem, com rigor metodológico adequado, de que forma recursos gamificados podem ser integrados a plataformas de saúde mental digital de maneira clinicamente embasada, considerando tipo de mecânica, perfil do usuário e correlação com estados emocionais registrados pelo sistema.

4.4 Sistema de reconhecimento emocional por voz

O modelo de reconhecimento emocional implementado neste trabalho constitui uma das camadas tecnológicas da plataforma. Os testes internos realizados demonstraram viabilidade técnica preliminar na classificação de emoções, conforme análise da matriz de confusão apresentada anteriormente.

Entretanto, a inteligência artificial não é tratada como elemento central isolado, mas como componente complementar dentro de um ecossistema maior de apoio emocional. O sistema permite a classificação probabilística da emoção detectada na fala, o registro histórico das variações emocionais e o apoio ao monitoramento longitudinal do usuário.

Ressalta-se que os testes foram conduzidos em ambiente controlado, com finalidade acadêmica e experimental, podendo os resultados variar conforme ampliação da base de dados, ajustes de hiperparâmetros e validação em contextos reais.

4.5 Integração dos recursos e potencial preventivo

O principal resultado do trabalho não reside exclusivamente na utilização de inteligência artificial, mas na integração sistêmica entre monitoramento automatizado, diário emocional, jogos interativos, possibilidade de anonimato e registro longitudinal de dados emocionais.

Essa integração permite que a plataforma atue em três níveis complementares:

- a) nível de expressão emocional, por meio do diário e dos desabafos;
- b) nível de regulação emocional, por meio dos jogos interativos;
- c) nível de monitoramento emocional, por meio da análise da fala e do histórico de registros.

Dessa forma, a solução proposta apresenta caráter multidimensional, diferenciando-se de abordagens baseadas exclusivamente em teleconsulta ou em análise automatizada isolada.

4.6 Impacto potencial e inovação

Os resultados indicam que a plataforma desenvolvida possui potencial inovador ao combinar tecnologia de reconhecimento emocional, ferramentas de expressão livre, estratégias lúdicas de enfrentamento, monitoramento contínuo e opção de anonimato controlado.

Essa combinação favorece maior engajamento do usuário e amplia as possibilidades de acompanhamento preventivo, contribuindo para a promoção do autocuidado emocional. Reforça-se que a proposta não objetiva substituir o atendimento psicológico profissional, mas atuar como ferramenta complementar de suporte e monitoramento.

4.7 Limitações e considerações éticas

Apesar dos resultados promissores, o trabalho apresenta limitações, dentre as quais destacam-se: realização de testes em ambiente controlado, ausência de validação em larga escala, dependência da qualidade dos dados de entrada e necessidade de ampliação da base de dados em língua portuguesa.

No que se refere aos aspectos éticos, ressalta-se a necessidade de consentimento informado, garantia de anonimato quando selecionado pelo usuário, proteção de dados sensíveis e transparência quanto à natureza probabilística das análises emocionais realizadas pelo sistema.

5 CONCLUSÃO

Os resultados obtidos demonstram que a plataforma proposta integra múltiplas estratégias de apoio emocional, combinando recursos tecnológicos e ferramentas interativas de maneira complementar. Verificou-se que o diário digital e a possibilidade de desabafo anônimo ampliam a acessibilidade emocional, enquanto os jogos interativos contribuem para regulação emocional e aumento do engajamento do usuário.

O reconhecimento emocional por voz apresentou viabilidade técnica preliminar, atuando como componente complementar dentro da arquitetura do sistema. Conclui-se que a inovação do trabalho reside na construção de um ecossistema digital de suporte emocional, no qual diferentes ferramentas operam de forma integrada, promovendo monitoramento, expressão e enfrentamento emocional de maneira estruturada e preventiva.

REFERÊNCIAS

FLEMING, Theresa M. et al. Serious games and gamification for mental health: current status and promising directions. *Frontiers in Psychiatry*, v. 7, p. 215, 2017. Disponível em: <https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsy.2016.00215/full>. Acesso em: 12 mar. 2026.

INKSTER, Becky; SARDA, Shyam; SUBRAMANIAN, Vinod. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. *JMIR mHealth and uHealth*, v. 6, n. 11, e12106, 2018. Disponível em: <https://mhealth.jmir.org/2018/11/e12106>. Acesso em: 12 mar. 2026.

GLOBO. IA Emocional: capacidade da Inteligência Artificial de compreender emoções humanas pode mudar para sempre as relações sociais; entenda. *O Globo*, 11 abr. 2025. Disponível em: <https://oglobo.globo.com/economia/tecnologia/noticia/2025/04/11/ia-emocional-capacidade-da-inteligencia-artificial-de-compreender-emocoes-humanas-pode-mudar-para-sempre-as-relacoes-sociais-entenda.ghtml>. Acesso em: 23 de nov de 2025.

LIVINGSTONE, Steven R.; RUSSO, Frank A. The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): a dynamic, multimodal set of facial and vocal expressions in North American English. *PLoS ONE*, v. 13, n. 5, e0196391, 2018. Disponível em: <https://doi.org/10.1371/journal.pone.0196391>. Acesso em: 12 mar. 2026.

Ministério da Saúde. Portaria GM/MS nº 1.836, de 27 de junho de 2022. Institui a Estratégia Nacional de Fortalecimento dos Cuidados da Ansiedade e Depressão no âmbito do SUS. Brasília, DF: Ministério da Saúde, 2022. Disponível em: https://bvsmis.saude.gov.br/bvs/saudelegis/doc_tec/27.06.2022_%20PRT_GM_1836.2022.pdf. Acesso em: 23 de novembro de 2025.

ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE. Pandemia de COVID-19 desencadeia aumento de 25% na prevalência de ansiedade e depressão no mundo. Brasília, DF: OPAS/OMS Brasil, 2022. Disponível em:

<https://www.paho.org/pt/noticias/2-3-2022-pandemia-covid-19-desencadeia-aumento-25-na-prevalencia-ansiedade-e-depressao-em>. Acesso em: 12 mar. 2026.

SAÚDE DIGITAL BRASIL. Cresce o número de atendimentos e a diversificação de produtos digitais para saúde mental e emocional entre associados da Saúde Digital Brasil. São Paulo: Saúde Digital Brasil, 2024. Disponível em: <https://saudedigitalbrasil.com.br/cresce-o-numero-de-atendimentos-e-a-diversificacao-de-produtos-digitais-para-saude-mental-e-emocional-entre-associados-da-saude-digital-brasil/>. Acesso em: 23 de novembro de 2025.

UNIVERSIDADE FEDERAL DO AMAZONAS (UFAM). Chatbot terapêutico utiliza inteligência artificial para atendimento emocional e psicológico. Portal UFAM – Notícias, 19 nov. 2024. Disponível em: <https://www.ufam.edu.br/noticias-destaque/6257-chatbot-terapeutico-utiliza-inteligencia-artificial-para-atendimento-emocional-e-psicologico.html>. Acesso em: 23 de novembro de 2025.

UNESP. Peixoto, G. S.; Linhares, J. E. B. Reconhecimento de emoções através da fala utilizando redes neurais convolucionais com ruído. Semish, 2023. Disponível em: <https://sol.sbc.org.br/index.php/semish/article/view/25067>. Acesso em: 23 de novembro de 2025.

UFPB. Aguiar, J. V. A. Reconhecimento de emoções na fala em português usando CNNs e LSTMs (TCC). João Pessoa: Universidade Federal da Paraíba, 2024. Disponível em: https://repositorio.ufpb.br/jspui/bitstream/123456789/34687/1/Joelder_Victor_Antonino_Aguiar_TCC.pdf. Acesso em: 23 de novembro de 2025.

VITTUDE. Tela inicial da plataforma Vittude. São Paulo, 2024. Disponível em: <https://vittude.com>. Acesso em: 24 de fevereiro de 2026.

WYSA. Wysa: everyday mental health. Londres: Wysa, 2016. Disponível em: <https://www.wysa.com>. Acesso em: 12 mar. 2026.

ZENKLUB. Tela inicial da plataforma Zenklub. São Paulo, 2024. Disponível em: <https://zenklub.com.br>. Acesso em: 24 de fevereiro de 2026.