**ANÁLISE COMPARATIVA DE ALGORITMOS DE SUPRESSÃO DE RUÍDO EM PLATAFORMAS DE COMUNICAÇÃO POR VOZ.**

**EDILSON PEREIRA DOS REIS SOUZA JUNIOR1**

**FABIO CASAGRANDE2**

**JOÃO VITOR DA COSTA VIEIRA3**

**MILTON TOMOMI OZEKI4**

**LUCIANO GONÇALVES DE CARVALHO 5**

**RESUMO**

Este trabalho analisa a eficácia dos algoritmos de supressão de ruído de seis plataformas populares de comunicação por voz: Discord (supressor padrão e Krisp), Microsoft Teams, Telegram, WhatsApp e Zoom. Utilizando gravações com diferentes tipos de ruído sobrepostos à fala, foram aplicadas análises visuais (espectrogramas) e quantitativas (similaridade espectral e nível médio em dBFS). Os resultados mostram que o Zoom e o Discord com Krisp apresentaram melhor equilíbrio entre remoção de ruído e preservação da voz, enquanto outras plataformas aplicaram filtragens menos eficazes.

# **Palavras-chave**: supressão de ruído, espectrograma, comunicação por voz, filtragem de áudio, plataformas digitais

# ABSTRACT

This study evaluates the noise suppression algorithms of six popular voice communication platforms: Discord (standard and Krisp), Microsoft Teams, Telegram, WhatsApp, and Zoom. Using recordings with various noise types mixed with speech, we conducted visual (spectrogram) and quantitative (spectral similarity and average dBFS level) analyses. The results indicate that Zoom and Discord with Krisp performed best in balancing noise reduction and voice clarity, while other platforms applied less efficient filters.

**Key words**: noise suppression, spectrogram, voice communication, audio filtering, digital platforms

# 1 INTRODUÇÃO

Em meio ao avanço da globalização, alimentada pelo desenvolvimento tecnológico dos meios de comunicação, a necessidade de uma maior flexibilidade e velocidade nos processos organizacionais utilizados por instituições em todo mundo criou uma demanda por métodos eficazes de comunicação que sejam capazes de conectar inúmeras pessoas distantes não só de forma rápida, mas instantânea. Entre as novas ferramentas que surgiram nesse cenário, as videoconferências se destacaram como um dos melhores meios disponíveis para suprir essa necessidade, uma vez que, através de vídeo e áudio, conseguem reunir pessoas de todos os lugares conectados à internet como se estivessem em uma mesma sala.

A crise sanitária do vírus SARS-CoV-2 em 2020, ao impedir o contato direto num mesmo ambiente, tornou o uso das videoconferências em uma necessidade diária em qualquer ambiente de trabalho. Segundo Tsuji Junior et al. (2023), o uso da videoconferência no contexto pandêmico evidenciou-se como uma alternativa viável e eficiente, com potencial de permanência mesmo após o período de crise sanitária.

Mesmo após a pandemia, os usos contínuos de sistemas de videoconferência perduraram, mostrando que os usos dessas ferramentas não estavam meramente atrelados a esses eventos, apesar de serem popularizados por eles. Hoje, reuniões online são necessárias e executadas nos mais diversos ambientes em que os conferentes estão sujeitos.

A qualidade do áudio se tornou um fator essencial para uma comunicação nos ambientes digitais. No entanto, diferentes fatores podem comprometer a clareza e a compreensão das transmissões de voz, principalmente os ruídos de fundo, interferências externas ou ecos. Esses problemas se tornam mais agravantes quando os participantes da reunião em questão se encontram em locais barulhentos, como escritórios compartilhados, cafés, ou mesmo em casa, onde sons indesejados podem prejudicar a compreensão da fala. Para mitigar esse problema, diversos algoritmos de supressão de ruídos foram desenvolvidos, buscando melhorar a experiência do usuário ao reduzir ou anular interferências indesejadas sem comprometer a inteligibilidade da voz.

Diante do cenário apresentado, o presente artigo visa documentar uma análise comparativa dos algoritmos de supressão de ruídos usados pelas plataformas: Whatsapp, Telegram, Zoom, Microsoft Teams e Discord. Avaliando seu desempenho em diferentes contextos com o intuito de demonstrar as vantagens e limitações de cada abordagem, contribuindo para a escolha de soluções mais eficazes na redução de interferências sonoras em comunicações digitais.

# 2 REFERENCIAL TEÓRICO

O ruído no senso comum é qualquer barulho indesejado, já para eletrônica qualquer corrente elétrica ou tensão indesejada observado na medição de um sinal elétrico (SALLES, 2017). Em reuniões remotas, os ruídos podem dificultar a compreensão, principalmente em ambientes de alta poluição sonora ou quando se usa equipamentos de baixa qualidade. Para mitigar esse problema, técnicas de supressão de ruído foram desenvolvidas para reduzir sons indesejados sem comprometer a voz do falante.

A supressão de ruídos é essencial em chamadas telefônicas, videoconferências e gravações de áudio no geral. Grandes plataformas como Google Meet e Discord utilizam algoritmos avançados para filtrar sons como ventiladores e latidos, melhorando a experiência do usuário. Além disso, softwares de edição de áudio permitem remover ruídos de gravações, alcançando maior nitidez na gravação.

Os métodos de supressão de ruídos envolvem técnicas de processamento digital de sinais (DSP), como filtragem de frequências e separação de fontes sonoras. A Transformada de Fourier, por exemplo, permite converter o áudio para o domínio da frequência, facilitando a remoção seletiva de ruídos (EMBARCADOS, 2023a; EMBARCADOS, 2023b). Modelos de aprendizado de máquina (machine learning) têm sido aplicados com sucesso na diferenciação entre voz humana e ruído de fundo, proporcionando alta precisão na limpeza de sinais sonoros (EMBARCADOS, 2023b).

Diferentes algoritmos são empregados na supressão de ruídos. Filtros adaptativos, como o Filtro de Wiener, ajustam dinamicamente a filtragem com base no som captado, proporcionando melhorias significativas na qualidade do áudio (ESTATÍSTICA FÁCIL, 2023). Outro exemplo é o algoritmo LMS (Least Mean Squares), que também realiza ajustes contínuos com base nos sinais recebidos (TELECO, 2023). Entre as técnicas que operam no domínio da frequência, destaca-se o uso de espectrogramas, que possibilitam a identificação visual de padrões sonoros e ruídos, auxiliando em sua supressão de forma seletiva.

O espectrograma, nesse contexto, é uma representação visual que mostra como o conteúdo de frequência de um sinal varia ao longo do tempo. Para construí-lo, o sinal é segmentado em pequenas janelas, e em cada uma delas é aplicada a Transformada de Fourier de curto termo (STFT), gerando o espectro de frequência instantâneo. Esses espectros são então dispostos em sequência temporal, criando uma visualização bidimensional em que a intensidade das cores ou níveis de brilho representa a amplitude das frequências em cada instante. Essa técnica permite identificar elementos como harmônicos, ruídos contínuos e sons transitórios, revelando características do áudio que não são perceptíveis apenas pela análise temporal (KHOZHAYEV, 2024).

Cada abordagem tem vantagens e desvantagens. Algoritmos baseados em filtros adaptativos são eficientes e podem ser usados em tempo real, mas têm limitações na remoção de ruídos complexos. Métodos de análise espectral oferecem precisão, mas podem comprometer a qualidade da fala. Já soluções baseadas em inteligência artificial conseguem resultados superiores, mas demandam mais recursos computacionais.

O principal desafio dessa tecnologia é preservar a voz original e evitar possíveis distorções. Processos excessivos podem tornar o áudio artificial, enquanto métodos mais simples podem não eliminar ruídos completamente. Além disso, a latência no processamento é um fator crítico, especialmente em aplicações ao vivo (VoIP), onde atrasos na transmissão da voz podem comprometer o entendimento da conversa.

Outro obstáculo é a adaptação a diferentes tipos de ruído. Sons inesperados, como conversas paralelas e barulhos repentinos, são difíceis de filtrar com técnicas tradicionais. Apesar dessas dificuldades, a evolução da supressão de ruídos continua a transformar a comunicação digital. O equilíbrio entre eficiência e custo computacional define quais algoritmos são mais adequados para cada aplicação, garantindo chamadas mais claras, assistentes virtuais mais responsivos e melhor qualidade de áudio em diferentes contextos.

Além das técnicas de supressão de ruído, é fundamental compreender as ferramentas utilizadas no processamento de áudio digital. Uma delas é a unidade dBFS (decibels relative to full scale), amplamente empregada em sistemas digitais como PCM, na qual 0 dBFS representa o nível máximo possível do sinal. Todos os níveis de pico abaixo desse valor são expressos como números negativos, indicando a margem abaixo do nível máximo permitido. Valores intermediários, como −6 dBFS, correspondem a 50 % do nível máximo, sendo o parâmetro padrão para monitoramento de amplitude em sistemas de áudio digital (ROHDE & SCHWARZ, 2025).

No contexto deste trabalho, o nível médio do sinal, uma estimativa de quão alto está o som, é calculado em dBFS a partir do valor RMS, uma prática consolidada em engenharia de áudio para representar o nível perceptivo de som . O uso desse parâmetro possibilita uma comparação objetiva entre os níveis dos áudios base (com ruído) e filtrados, permitindo avaliar o impacto das técnicas de supressão.

Em termos de implementação, utilizamos a biblioteca Librosa, uma das mais robustas e confiáveis para processamento de sinais de áudio em Python. Criada originalmente para pesquisas em Music Information Retrieval (MIR), ela fornece funções para leitura de arquivos, cálculo de STFT, extração de recursos e visualização (McFEE et al., 2015). Por sua ampla adoção e documentação, Librosa se tornou referência em aplicações que envolvem análise espectral e manipulação de sinais de fala, como no presente estudo.

# 3 MATERIAL E MÉTODOS

A seleção dos algoritmos de filtragem de ruído analisados neste estudo foi baseada na popularidade global das plataformas que oferecem recursos de chamadas de voz. Foram considerados, especificamente, os aplicativos com o maior número estimado de usuários ativos mensais e com ampla adoção internacional, com dados coletados em 26 de maio de 2025.

Foram incluídos os seguintes aplicativos:

* WhatsApp (Meta): aproximadamente 3 bilhões de usuários ativos mensais;
* Telegram: cerca de 1 bilhão de usuários ativos mensais;
* Zoom: aproximadamente 300 milhões de usuários ativos mensais;
* Microsoft Teams: estimativa de 320 milhões de usuários ativos mensais, com cerca de 280 milhões ativos diariamente até 2023;
* Discord: aproximadamente 563 milhões de usuários ativos mensais, com base nas estimativas mais recentes disponíveis até 2023.

Esses dados foram obtidos a partir de fontes consolidadas de estatísticas globais de uso, principalmente da plataforma Statista, que fornece atualizações frequentes sobre o número de usuários ativos em aplicativos de comunicação e redes sociais [Statista, 2024a; Statista, 2024b; Statista, 2024c].

Note que estes não representam exatamente os cinco aplicativos mais utilizados globalmente, a escolha foi norteada por critérios de inclusão e exclusão específicos. Os critérios de exclusão aplicados foram:

* Ausência de recurso de chamada de voz;
* Popularidade geograficamente restrita, com uso concentrado em regiões específicas (por exemplo, aplicativos amplamente utilizados apenas na China), o que limita sua relevância para o contexto local da análise.
* Disponibilidade do recurso de chamadas de voz apenas mediante pagamento;
* Baixa notoriedade ou uso marginal do recurso de chamada de voz;
* Foco exclusivo ou predominante em chamadas informais, nas quais a supressão de ruído não representa um requisito relevante.

Aplicativos como YouTube, TikTok, Weibo, Pinterest e Twitch foram excluídos por não oferecerem chamadas de voz. Plataformas como WeChat, QQ, Qzone, Kuaishou, Tieba, Viber, imo e Line foram desconsideradas por sua baixa penetração ou relevância fora do contexto asiático. O Twitter foi excluído por restringir chamadas de voz a usuários pagantes, enquanto o LinkedIn e o Reddit foram desconsiderados por apresentarem funcionalidades pouco conhecidas ou subutilizadas. Aplicativos com foco informal, como o Snapchat, também não foram incluídos devido à menor importância atribuída à filtragem de ruído nesse tipo de comunicação.

Adicionalmente, os aplicativos WhatsApp, Messenger, Facebook e Instagram fazem parte do mesmo conglomerado empresarial, a Meta. Ainda que não haja confirmação oficial de que utilizam exatamente o mesmo algoritmo de supressão de ruído, é razoável supor que utilizem mecanismos semelhantes de filtragem baseados em aprendizado de máquina, com possíveis ajustes conforme o perfil de uso de cada aplicativo.

No caso do Discord, destaca-se a presença de dois algoritmos distintos de supressão de ruído disponíveis aos usuários: um algoritmo próprio da plataforma, e a tecnologia fornecida pela empresa Krisp, especializada em filtragem de ruído ambiente com uso de inteligência artificial. Essa dualidade implica na análise de ambos os algoritmos neste presente artigo.

Para avaliar a eficácia dos algoritmos de supressão de ruído das plataformas selecionadas, foi selecionado um áudio de teste composto por uma gravação de fala clara, uma locução com entonação natural, dicção nítida e ritmo moderado, representando uma situação típica de chamada de voz em ambiente cotidiano. Este áudio foi combinado com diferentes tipos de ruídos de fundo, com o propósito de testar os algoritmos e comparar o áudio sem ruído, com o ruído e com o ruído suprimido.

Foram selecionados cinco tipos de ruído ambiental comumente presentes em cenários reais de comunicação: choro de bebê, sons de obras, latidos de cachorro, passagem de caminhão e ruído de multidão barulhenta. Cada ruído foi sobreposto à gravação da fala, resultando em cinco amostras distintas de áudio contendo simultaneamente voz e interferência sonora. Cada uma dessas amostras foi processada separadamente por seis algoritmos de supressão de ruído, correspondentes às plataformas analisadas no estudo. Dessa forma, obtivemos 30 arquivos resultantes, possibilitando uma comparação ampla e sistemática do desempenho de cada tecnologia de filtragem em diferentes contextos de interferência sonora.

As amostras com ruído foram então processadas pelos algoritmos de supressão de cada uma das cinco plataformas analisadas. Para isso, o grupo realizou a filtragem dos áudios utilizando o software VoiceMeeter, configurado para criar um microfone virtual. As amostras foram reproduzidas por esse microfone virtual e capturadas conforme os recursos de gravação disponíveis em cada aplicação. No Discord, utilizou-se a função de teste de microfone presente nas configurações de voz e vídeo, permitindo ouvir como o áudio era processado pela plataforma, e gravando com o aplicativo OBS. No Microsoft Teams e no Zoom, foi iniciada uma gravação de reunião, funcionalidade nativa de ambas as plataformas, enquanto as amostras eram reproduzidas com o VoiceMeeter; em seguida, os áudios gravados foram baixados para análise. Já no Telegram e no WhatsApp, os áudios foram reproduzidos via microfone virtual e gravados diretamente através da funcionalidade de envio de mensagens de voz de cada aplicativo. Posteriormente, as gravações foram extraídas para avaliação. A comparação entre os sinais de entrada (áudio com ruído) e os sinais de saída (áudio pós-processamento) possibilita a avaliação qualitativa e quantitativa do desempenho dos algoritmos na atenuação ou remoção do ruído de fundo.

Para realizar a análise qualitativa das alterações sonoras provocadas pelos algoritmos de supressão de ruído, sobrepusemos visualmente os espectrogramas das amostras de áudio. A imagem do espectrograma do áudio base (sem filtragem) foi posicionada como camada inferior, enquanto o espectrograma da versão filtrada, gerada após o processamento por uma das plataformas, foi colocado sobre a anterior com transparência. Essa sobreposição permite verificar visualmente quais regiões espectrais foram atenuadas ou preservadas pelos algoritmos de filtragem. Para isso, desenvolvemos um script em Python utilizando a biblioteca Librosa para extração e visualização dos espectrogramas.

Além da análise visual, também com o auxílio de scripts em Python, realizamos uma avaliação quantitativa das amostras por meio do cálculo de métricas relevantes para o contexto de supressão de ruído:

* Similaridade Espectral entre o áudio filtrado com o áudio original sem ruído
* Nível Médio do Áudio Base em dBFS
* Nível Médio do Áudio Filtrado em dBFS

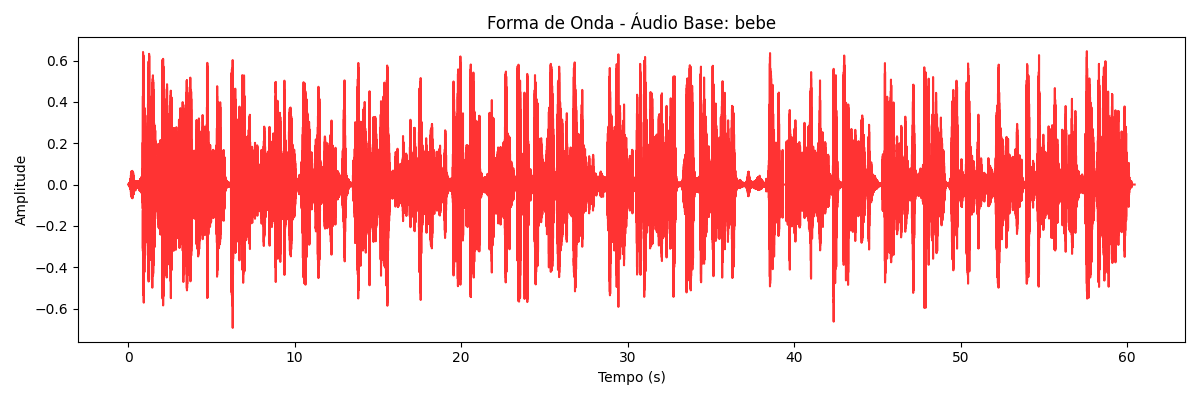
Essa metodologia busca reproduzir de forma controlada condições práticas de uso, nas quais a presença de ruído ambiental é frequente e/ou variável.

# 4 RESULTADOS E DISCUSSÃO

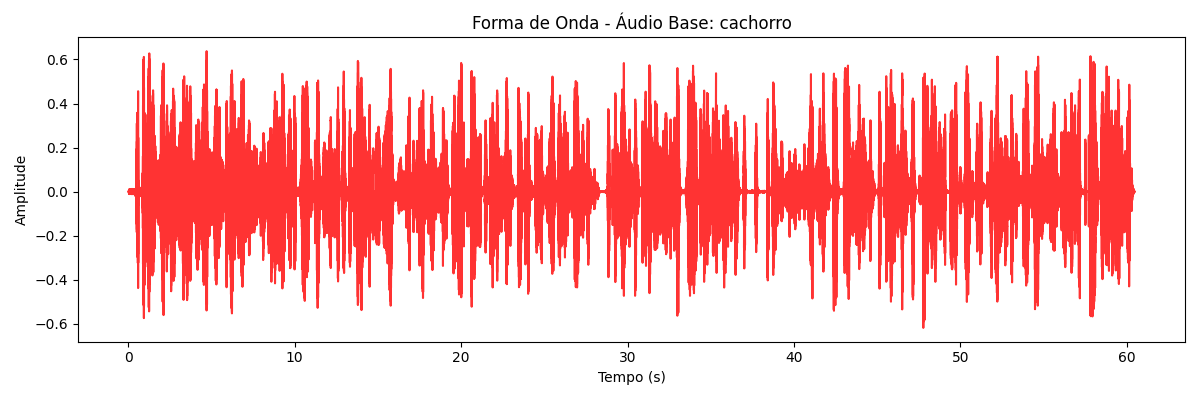
**4.1 Análise visual dos espectrogramas**

A seguir, apresentam-se os espectrogramas das amostras de áudio utilizadas nos testes, compostas por uma gravação de voz clara combinada com diferentes tipos de ruído de fundo, representando cenários cotidianos de comunicação. Cada figura ilustra a interação entre o sinal vocal e o respectivo ruído:

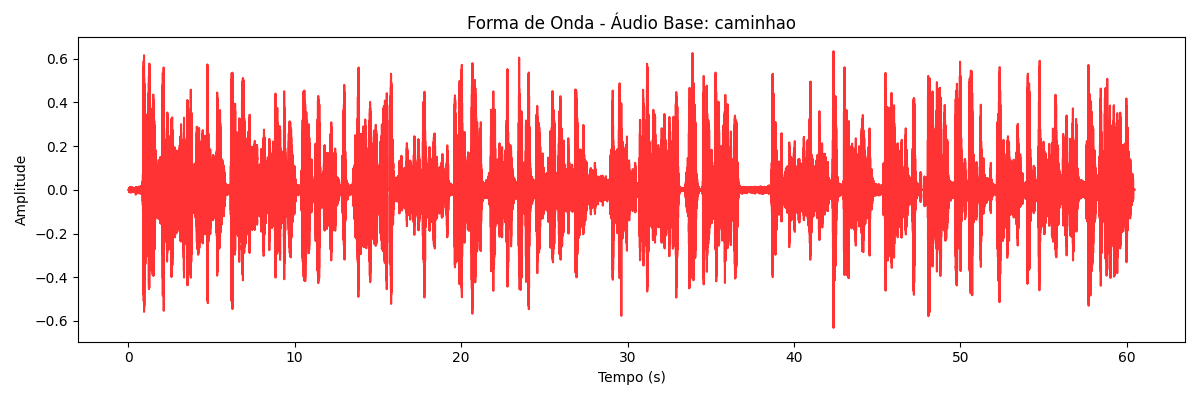
**Figura 1** – Voz com som de bebê chorando ao fundo



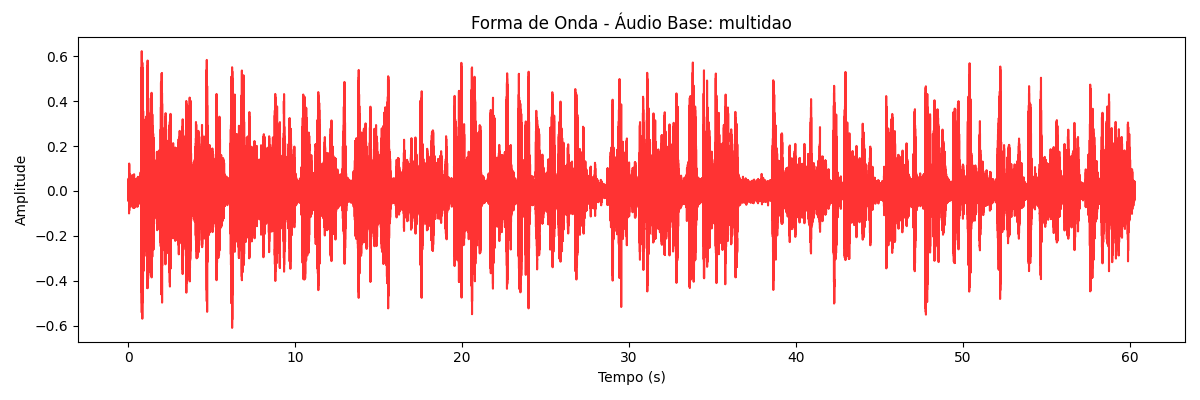
**Figura 2** – Voz com latidos de cachorro ao fundo



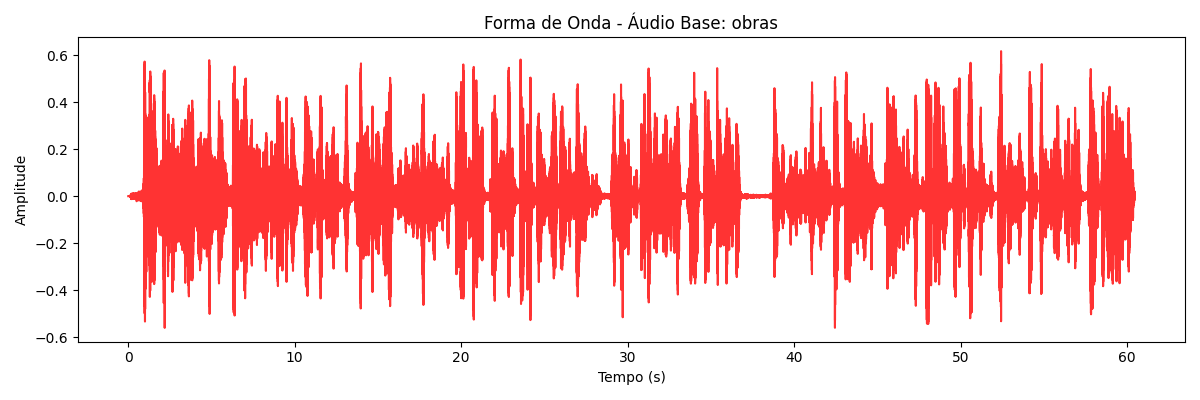
**Figura 3** – Voz com ruído de caminhão passando



**Figura 4** – Voz com ruído de multidão conversando



**Figura 5** – Voz com barulhos de construção civil

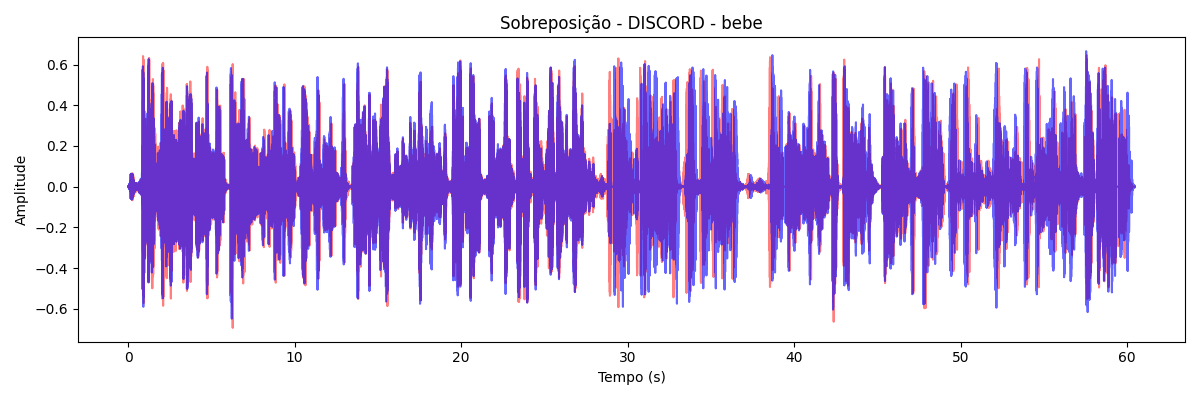


**Fonte**: Os autores.

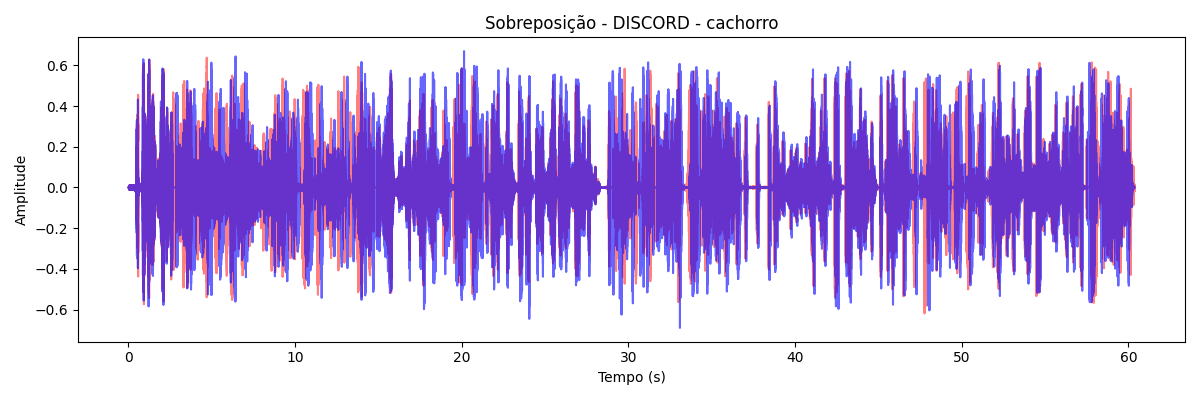
A seguir, apresentamos os espectrogramas sobrepostos, organizados por plataforma, referentes às cinco amostras de áudio utilizadas nos testes. Em cada figura, o espectrograma do áudio base (com ruído) está representado como camada inferior, enquanto o espectrograma da versão filtrada, resultante do algoritmo de supressão da plataforma, aparece sobreposto com transparência, permitindo uma análise visual direta das alterações espectrais provocadas pelo processamento.

**4.1.1 Discord (supressor padrão)**

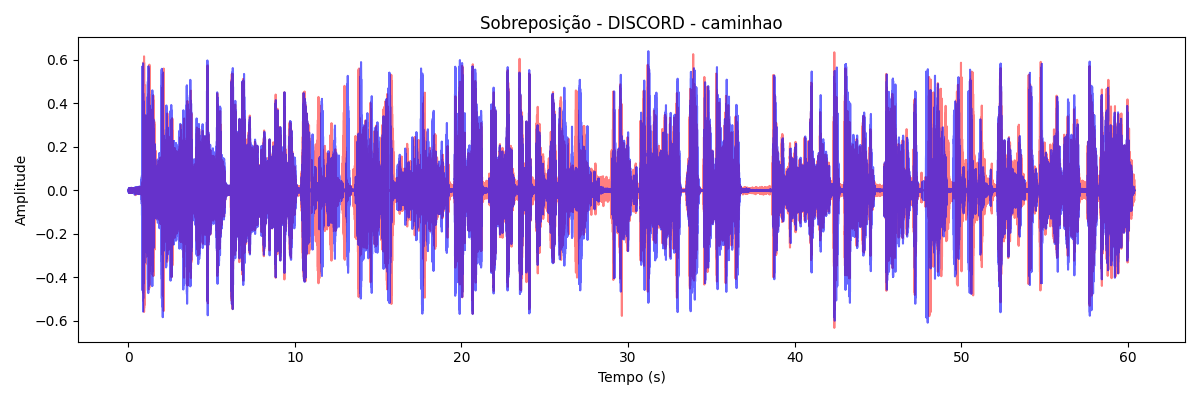
**Figura 6** – Espectrograma sobreposto: voz com som de bebê (Discord)



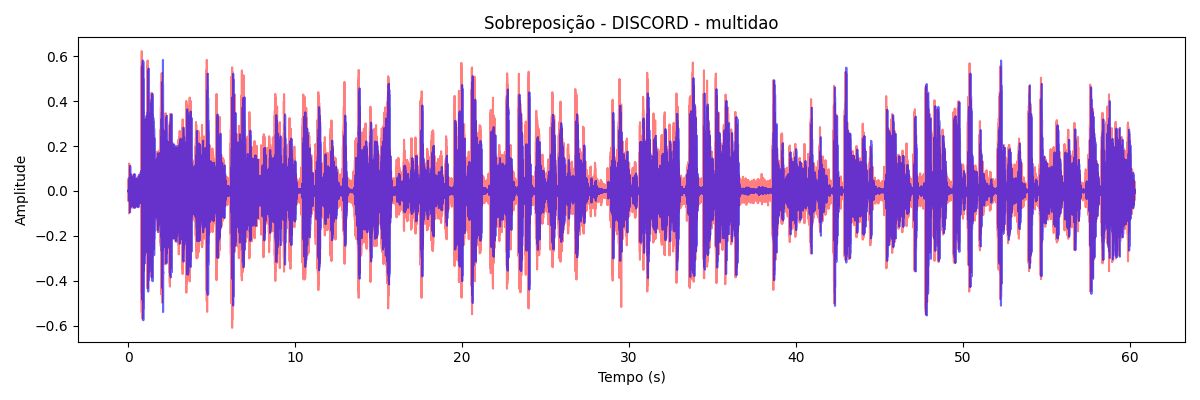
**Figura 7** – Espectrograma sobreposto: voz com latidos de cachorro (Discord)



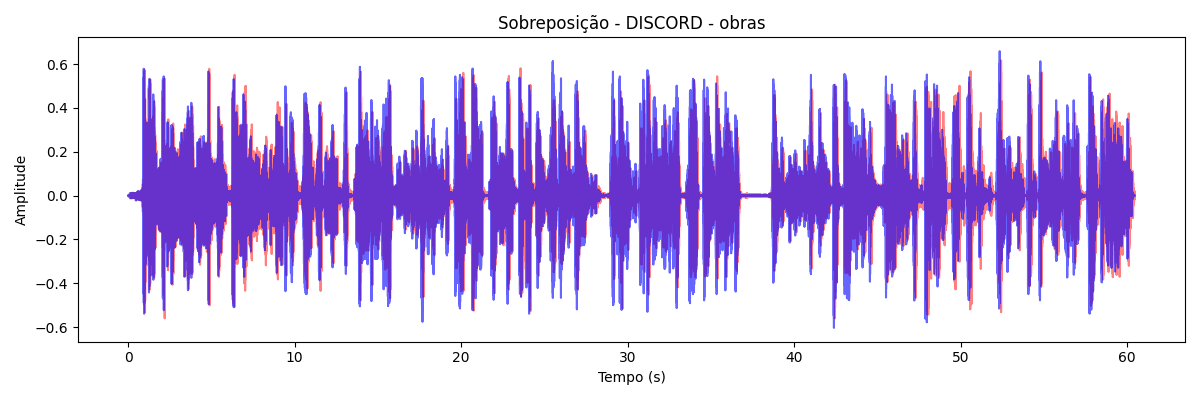
**Figura 8** – Espectrograma sobreposto: voz com ruído de caminhão (Discord)



**Figura 9** – Espectrograma sobreposto: voz com ruído de multidão (Discord)



**Figura 10** – Espectrograma sobreposto: voz com ruído de construção (Discord)

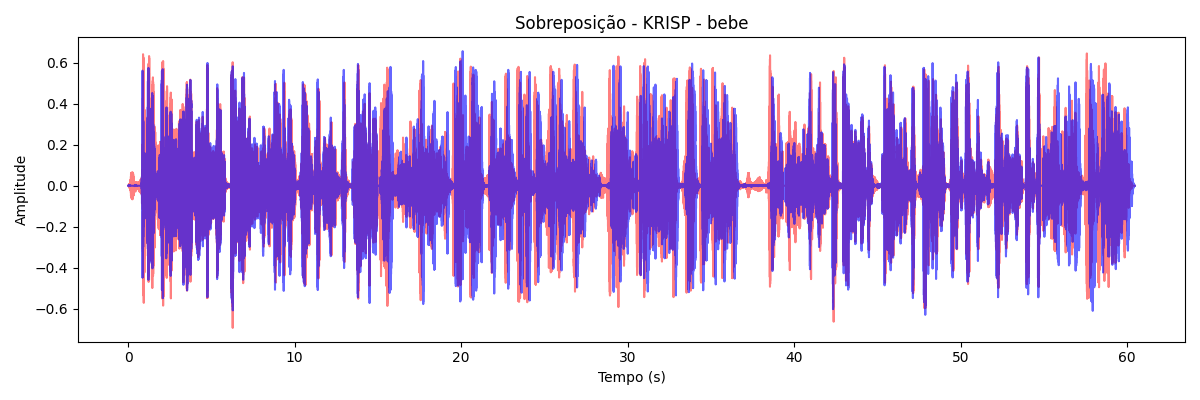


**Fonte**: Os autores.

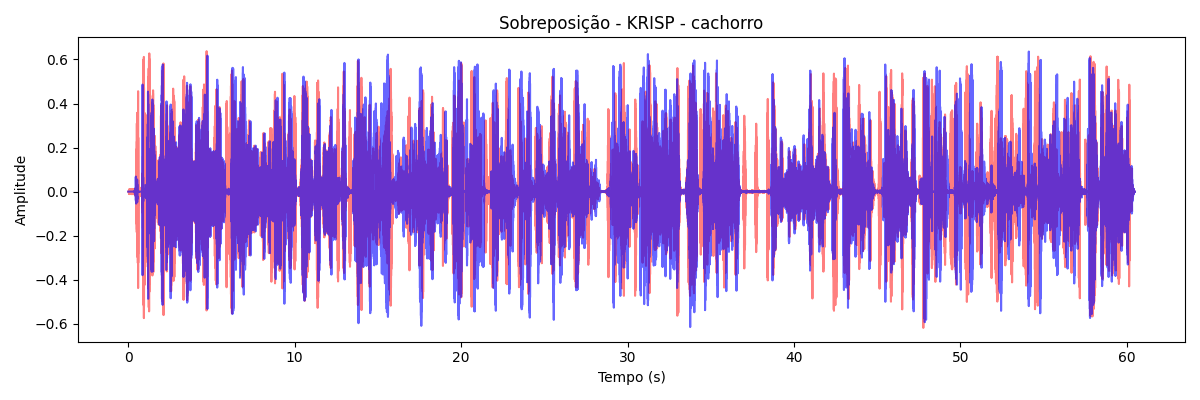
Na filtragem do Discord, com supressor padrão, foi possível perceber uma ineficiência em filtrar os ruídos dos áudios. Com exceção do ruído de multidão, os ruídos se apresentaram semelhantes aos áudios anteriores à filtragem.

**4.1.2 Discord (Krisp)**

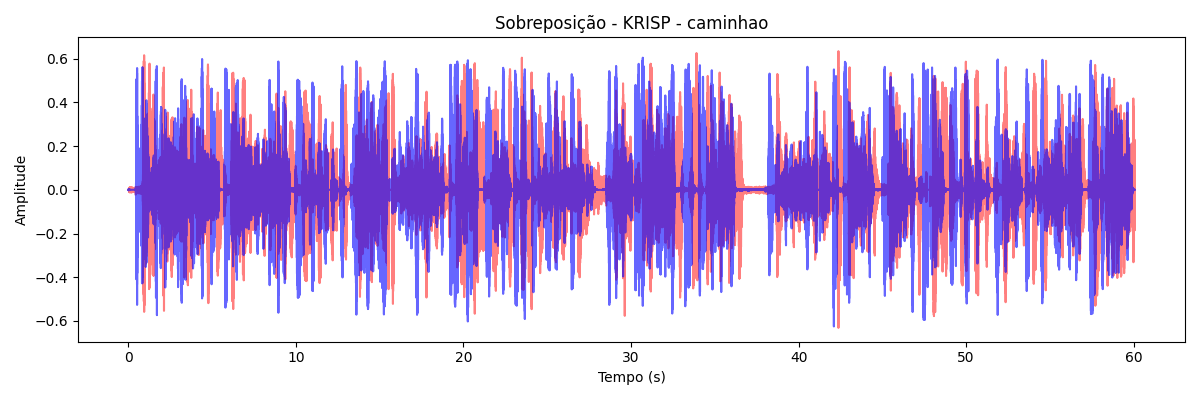
**Figura 11** – Espectrograma sobreposto: voz com som de bebê (Krisp)



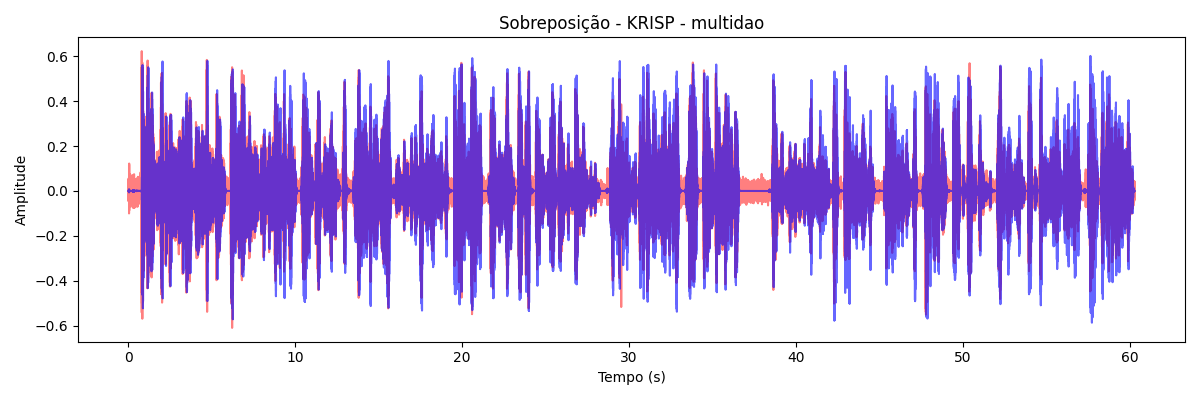
**Figura 12** – Espectrograma sobreposto: voz com latidos de cachorro (Krisp)



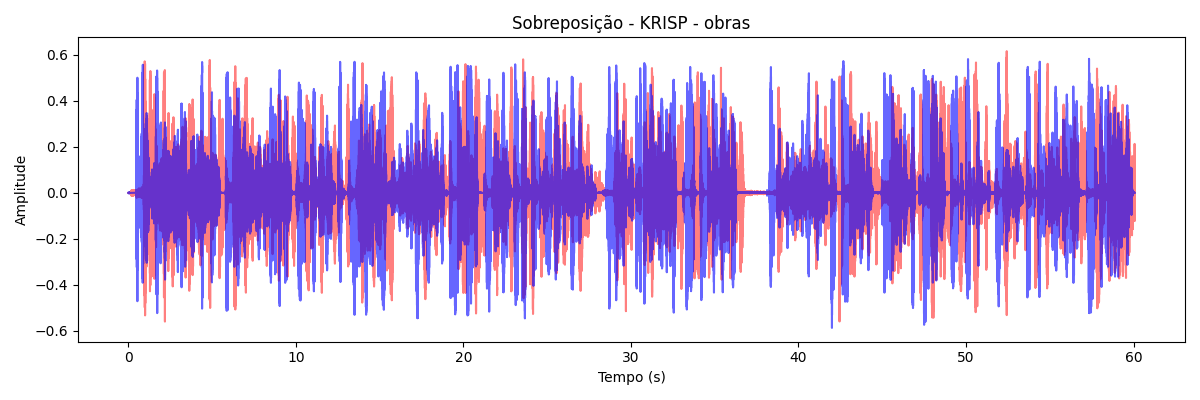
**Figura 13** – Espectrograma sobreposto: voz com ruído de caminhão (Krisp)



**Figura 14** – Espectrograma sobreposto: voz com ruído de multidão (Krisp)



**Figura 15** – Espectrograma sobreposto: voz com ruído de construção (Krisp)

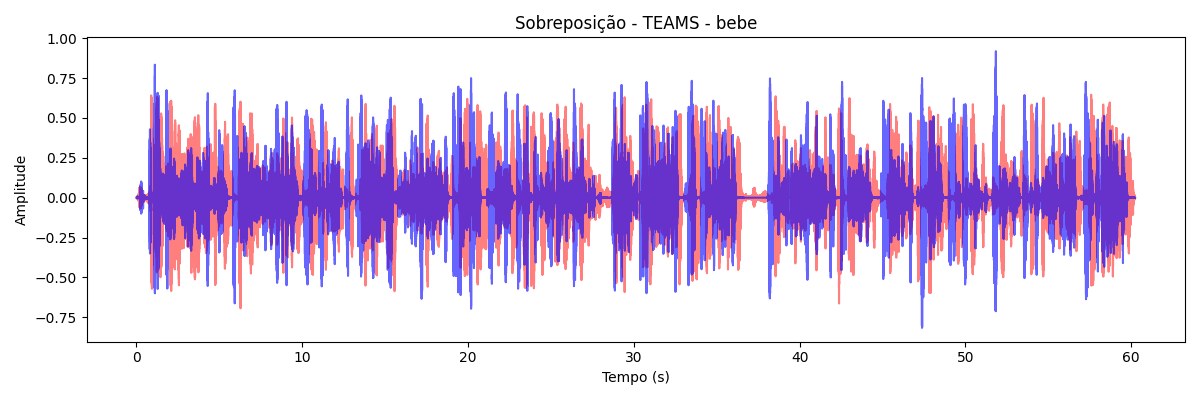


**Fonte**: Os autores.

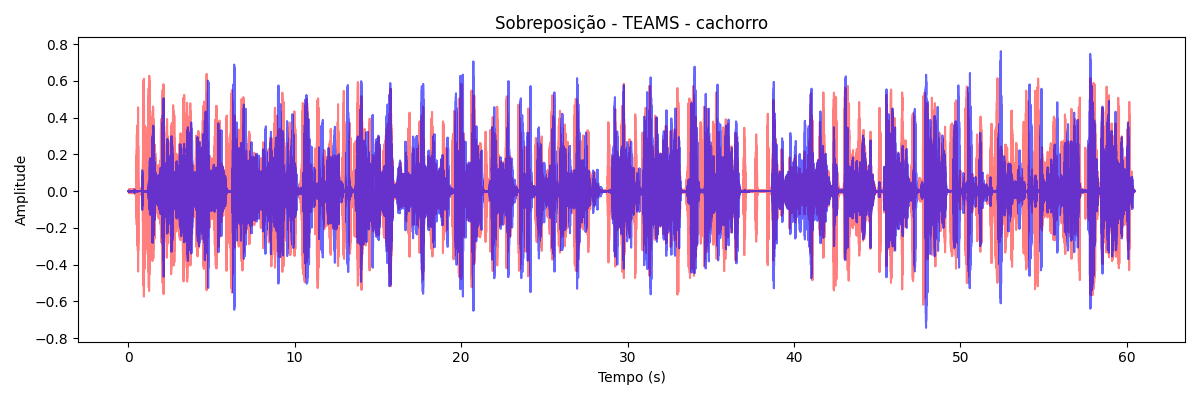
A filtragem de ruídos feita pelo algoritmo Krisp, no Discord, se mostrou muito mais eficiente, este foi capaz de suprimir grande parte dos ruídos. Mesmo que não tenha suprimido os ruídos completamente, este supressor possibilita uma comunicação clara entre as partes de uma conversa nas situações testadas.

**4.1.3 Teams**

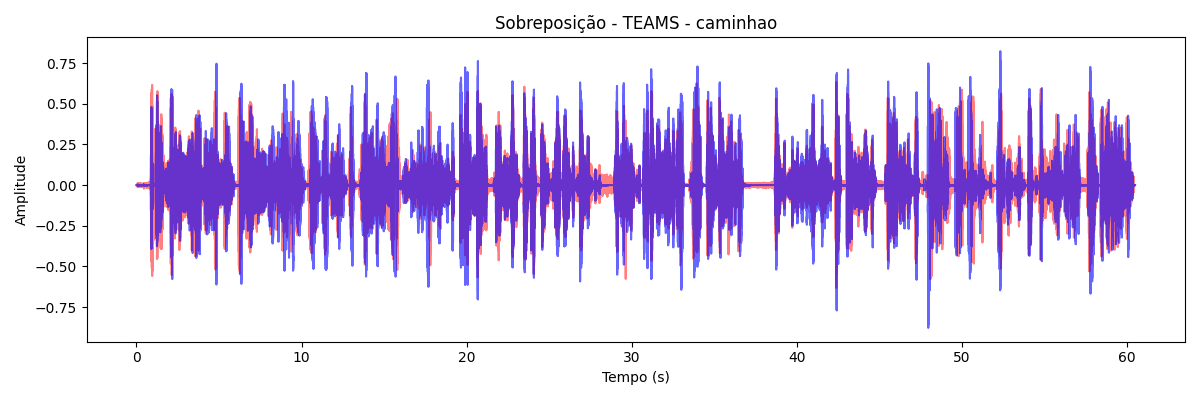
**Figura 16** – Espectrograma sobreposto: voz com som de bebê (Teams)



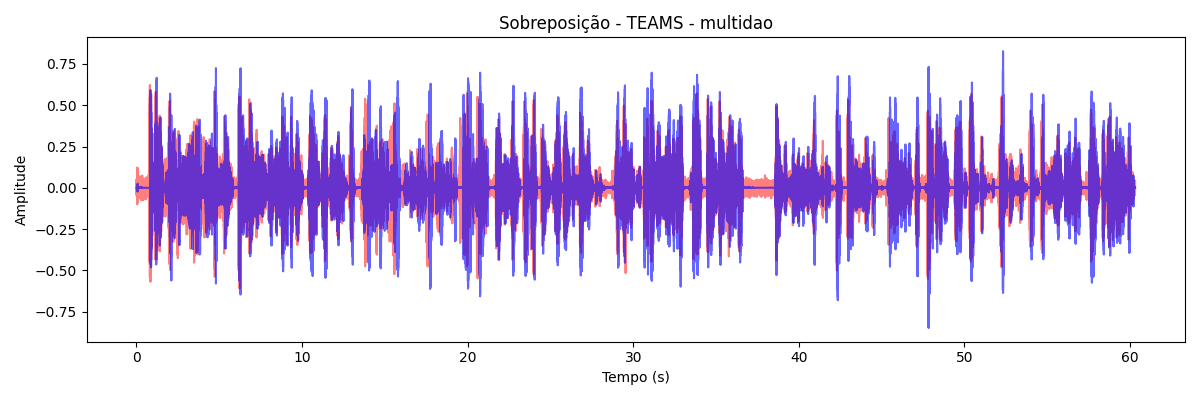
**Figura 17** – Espectrograma sobreposto: voz com latidos de cachorro (Teams)



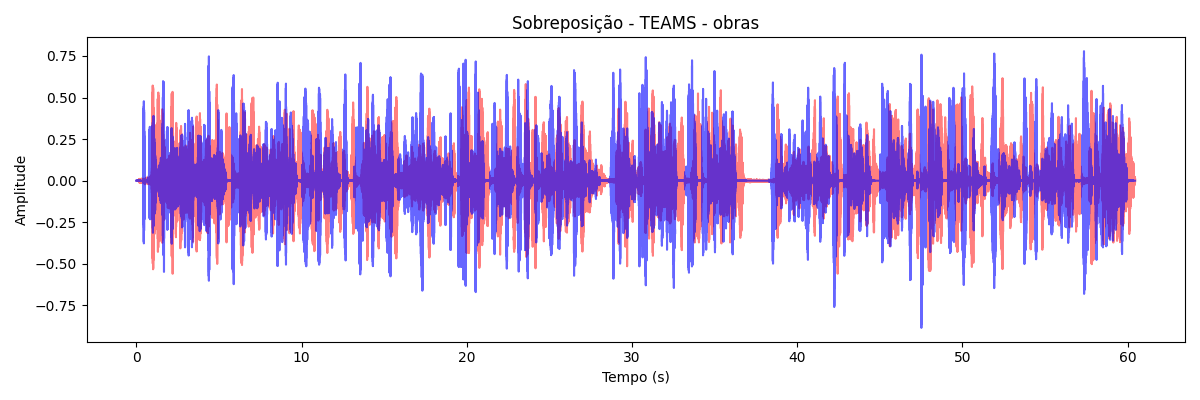
**Figura 18** – Espectrograma sobreposto: voz com ruído de caminhão (Teams)



**Figura 19** – Espectrograma sobreposto: voz com ruído de multidão (Teams)



**Figura 20** – Espectrograma sobreposto: voz com ruído de construção (Teams)

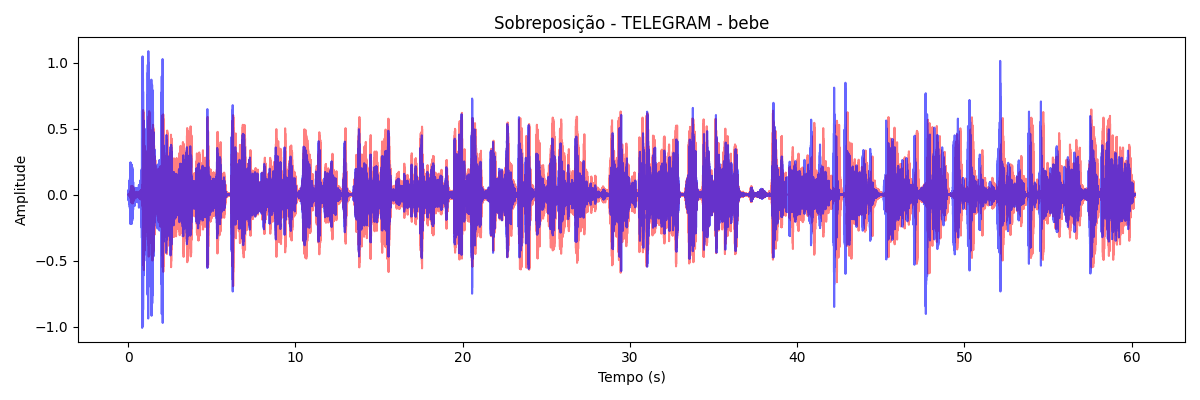


**Fonte**: Os autores.

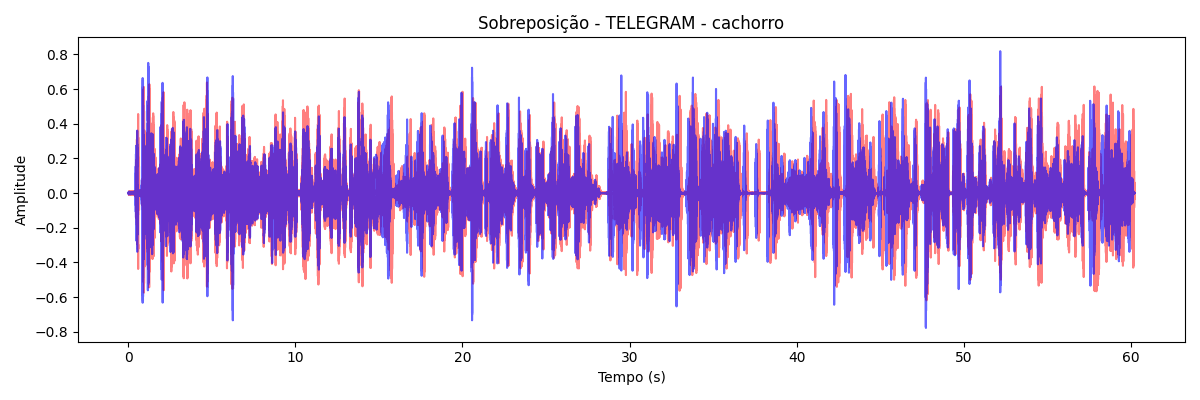
O algoritmo de supressão do Teams foi capaz de filtrar parte considerável dos ruídos. Não os eliminou completamente, mas diminui a intensidade, que, em uma situação de conversa, pode ser considerado satisfatório.

**4.1.4 Telegram**

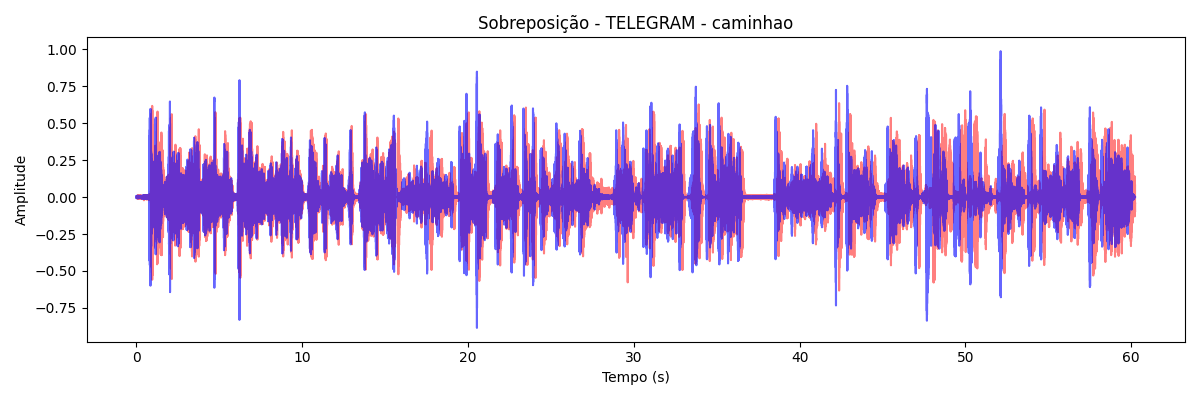
**Figura 21** – Espectrograma sobreposto: voz com som de bebê (Telegram)



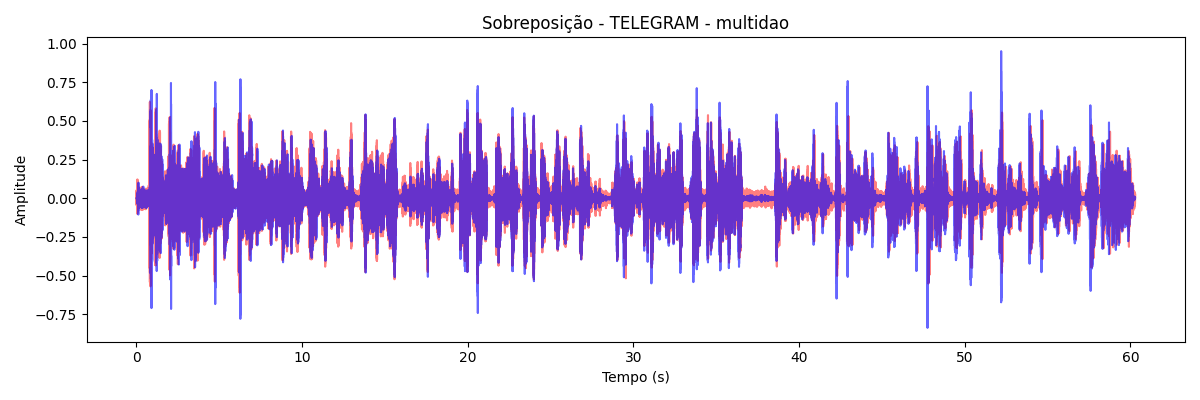
**Figura 22** – Espectrograma sobreposto: voz com latidos de cachorro (Telegram)



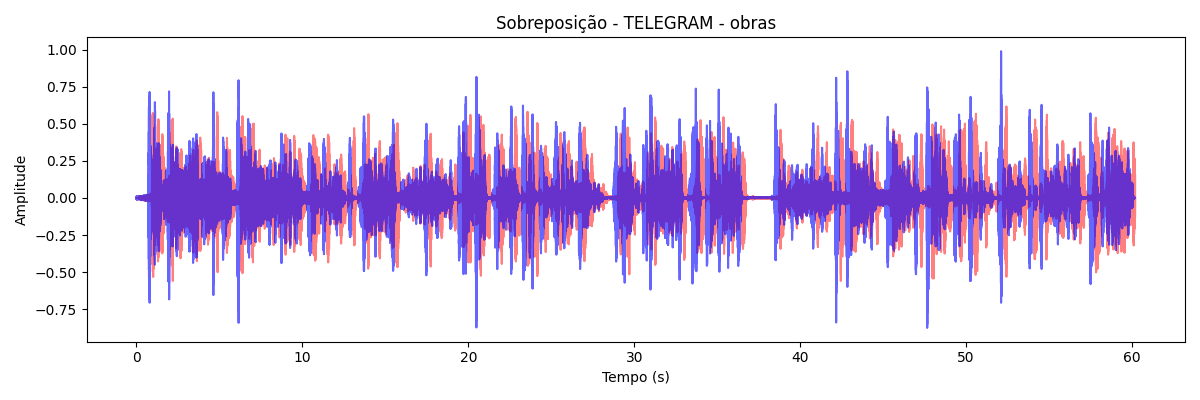
**Figura 23** – Espectrograma sobreposto: voz com ruído de caminhão (Telegram)



**Figura 24** – Espectrograma sobreposto: voz com ruído de multidão (Telegram)



**Figura 25** – Espectrograma sobreposto: voz com ruído de construção (Telegram)

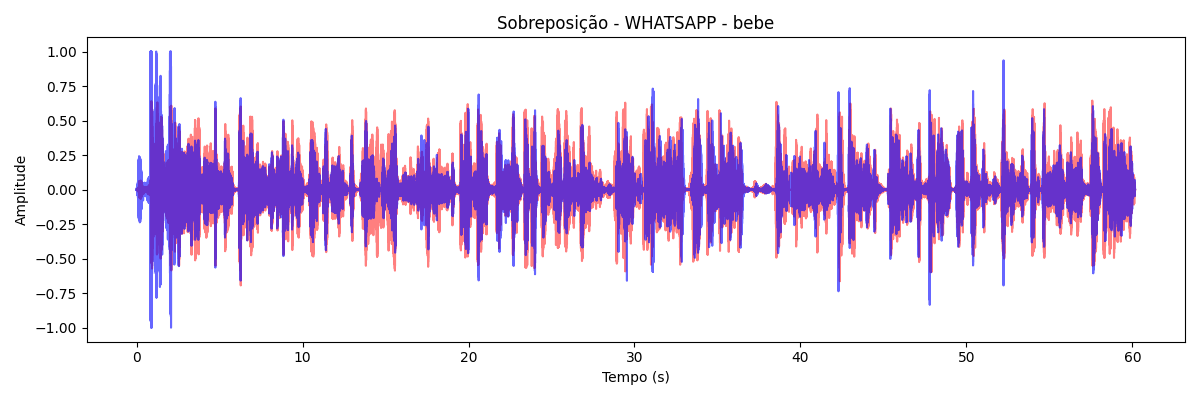


**Fonte**: Os autores.

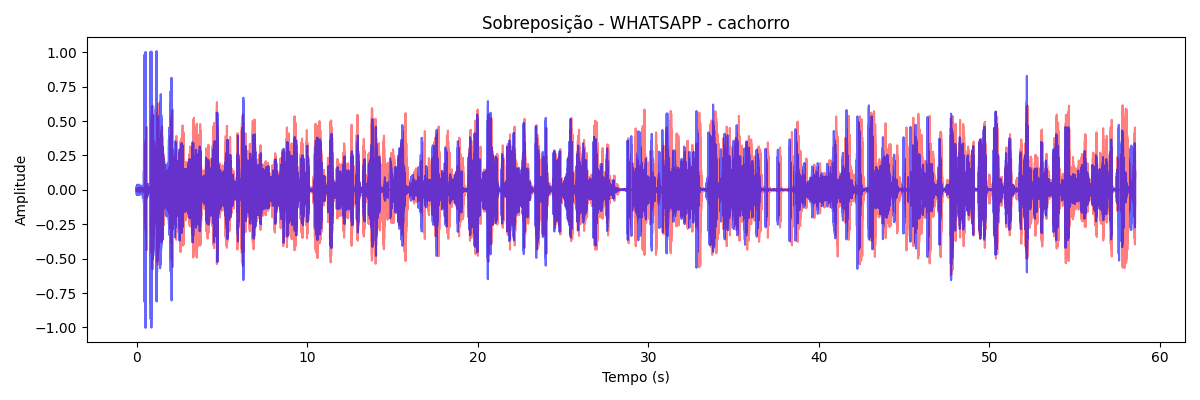
Os áudios filtrados pelo supressor de ruídos do Telegram apresentaram uma pequena melhora em relação às suas versões pré-processadas. Foi capaz de diminuir um pouco o volume do ruído, mas não removendo-o. Especialmente, apresentou dificuldade em filtrar o ruído de multidão.

**4.1.5 Whatsapp**

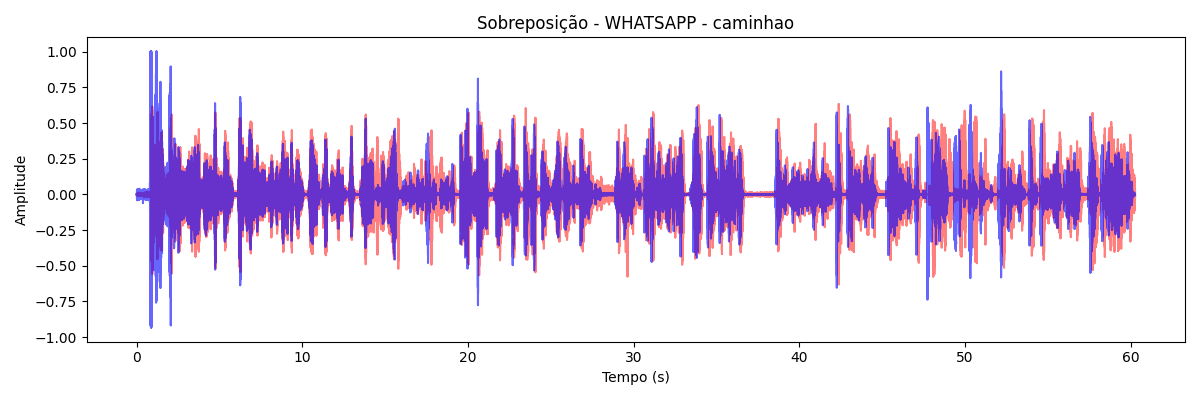
**Figura 26** – Espectrograma sobreposto: voz com som de bebê (Whatsapp)



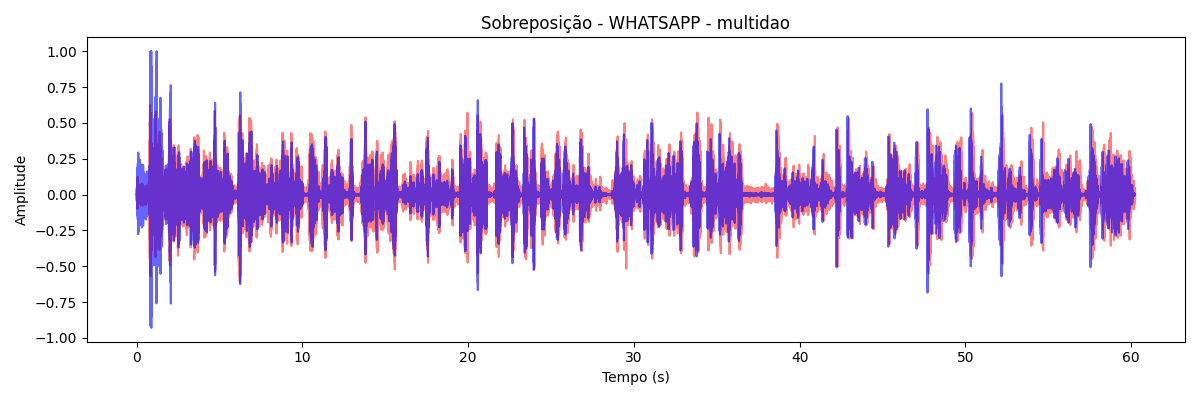
**Figura 27** – Espectrograma sobreposto: voz com latidos de cachorro (Whatsapp)



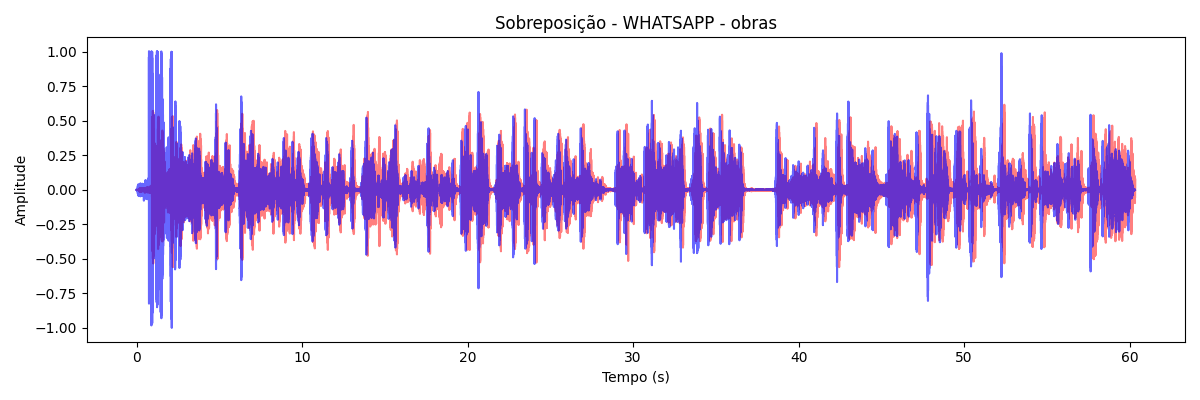
**Figura 28** – Espectrograma sobreposto: voz com ruído de caminhão (Whatsapp)



**Figura 29** – Espectrograma sobreposto: voz com ruído de multidão (Whatsapp)



**Figura 30** – Espectrograma sobreposto: voz com ruído de construção (Whatsapp)

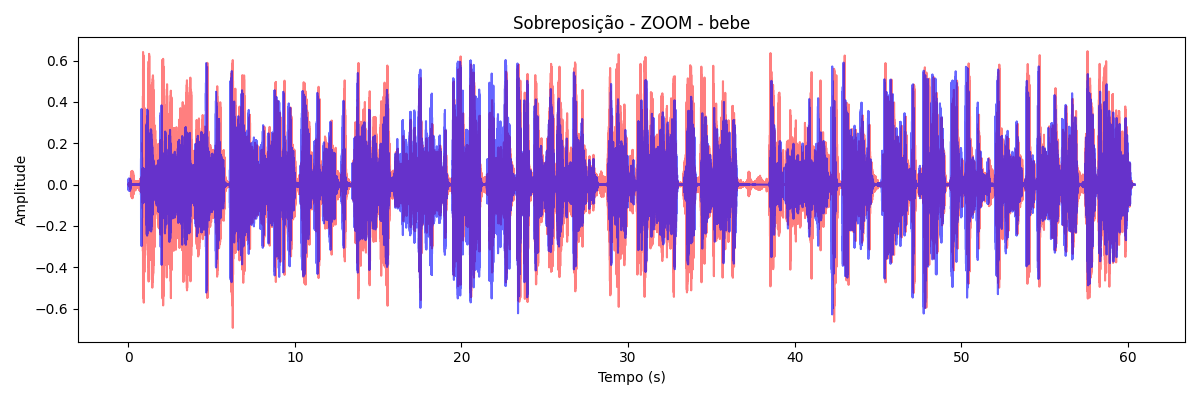


**Fonte**: Os autores.

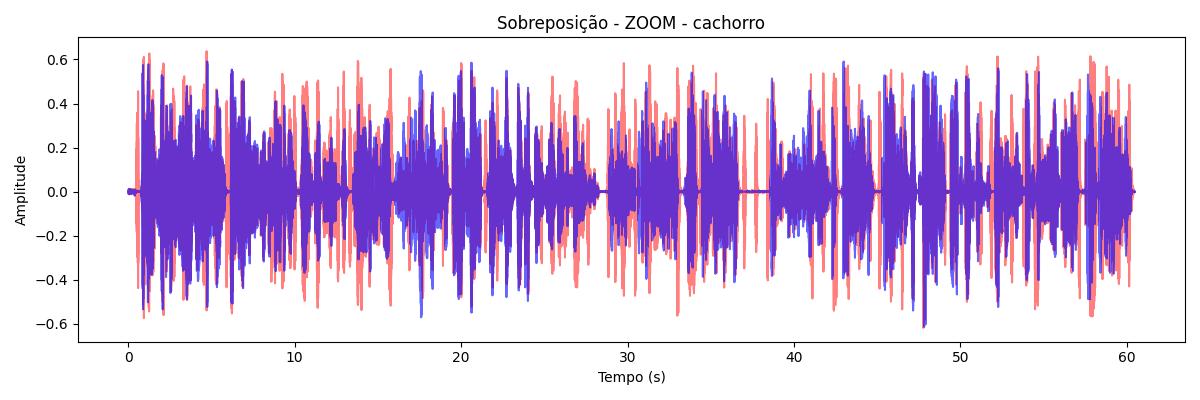
Na supressão de ruídos do Whatsapp, foi possível perceber que ele se mostrou incapaz de remover os ruídos em algumas situações. No áudio com ruído de caminhão, entretanto, ele se mostrou mais eficaz.

**4.1.6 Zoom**

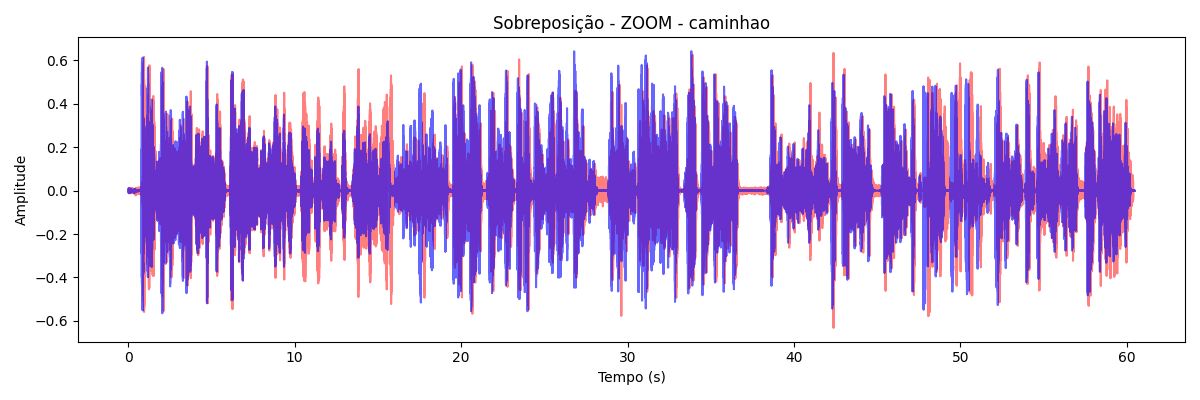
**Figura 31** – Espectrograma sobreposto: voz com som de bebê (Zoom)



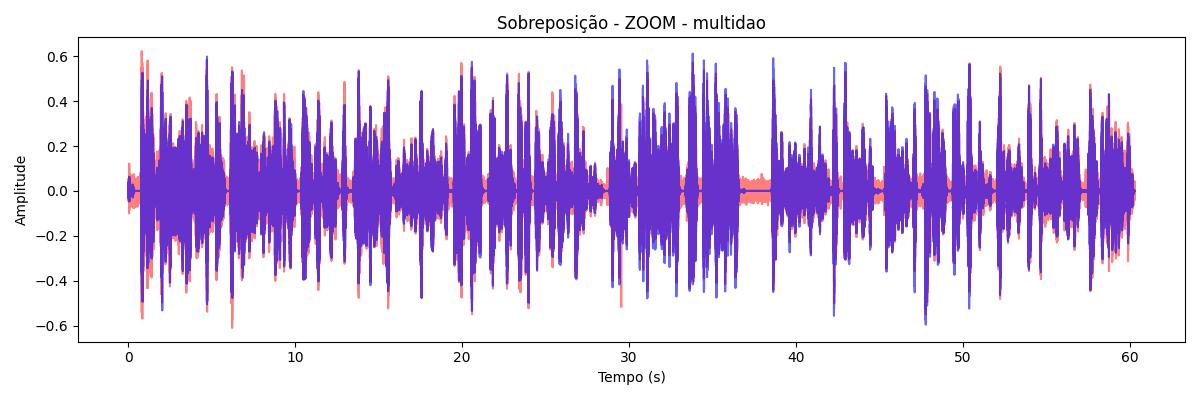
**Figura 32** – Espectrograma sobreposto: voz com latidos de cachorro (Zoom)



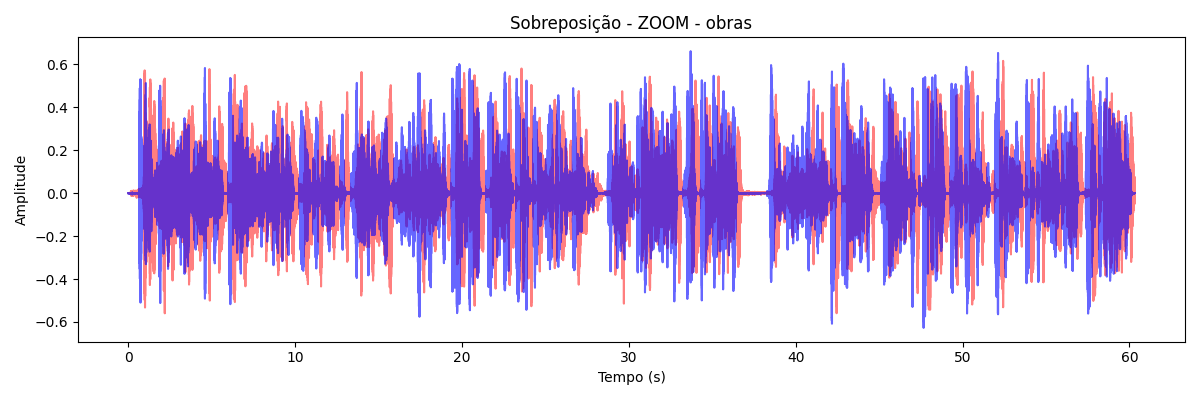
**Figura 33** – Espectrograma sobreposto: voz com ruído de caminhão (Zoom)



**Figura 34** – Espectrograma sobreposto: voz com ruído de multidão (Zoom)



**Figura 35** – Espectrograma sobreposto: voz com ruído de construção (Zoom)



**Fonte**: Os autores.

O algoritmo de filtragem de ruídos do Zoom conseguiu preservar bem a voz falante, ao passo que também conseguiu minar os ruídos do áudio. Com destaque especial para o ruído de multidão, onde o Zoom teve sua maior performance na filtragem de áudio.

**4.2 Métricas quantitativas**

Como citado anteriormente, foi calculado métricas quantitativas referentes às filtragens do áudio, a seguir serão apresentadas:

**4.2.1 Discord (supressor padrão)**

**Quadro 1** – Métricas do Discord (supressor padrão)

| **Ruído** | **Similaridade Espectral com Original (%)** | **Nível Médio (Base) [dBFS]** | **Nível Médio (Filtrado) [dBFS]** |
| --- | --- | --- | --- |
| bebe | 64,84% | -17,37 | -16,49 |
| cachorro | 53,52% | -17,01 | -16,56 |
| caminhão | 53,45% | -17,8 | -17,19 |
| multidão | 93,69% | -18,31 | -19,2 |
| obras | 97,72% | -18,04 | -17,51 |

**Fonte**: Os autores.

Os dados mostram que o Discord tem melhor desempenho com ruídos contínuos, como multidão e obras, alcançando mais de 93% de similaridade espectral com o áudio original. Já ruídos pontuais, como bebê e cachorro, reduzem essa similaridade para cerca de 53% a 65%. Isso indica que o supressor é mais eficaz com ruídos estáveis. Os níveis médios em dBFS mostram pouca perda de intensidade na voz.

**4.2.2 Discord (Krisp)**

**Quadro 2** – Métricas do Discord (Krisp)

| **Ruído** | **Similaridade Espectral com Original (%)** | **Nível Médio (Base) [dBFS]** | **Nível Médio (Filtrado) [dBFS]** |
| --- | --- | --- | --- |
| bebe | 58,64% | -17,37 | -16,75 |
| cachorro | 53,49% | -17,01 | -16,72 |
| caminhão | 54,08% | -17,8 | -16,47 |
| multidão | 96,75% | -18,31 | -16,69 |
| obras | 87,08% | -18,04 | -16,54 |

**Fonte**: Os autores.

O Discord com Krisp apresentou alta similaridade espectral para ruídos contínuos, como multidão (96,75%) e obras (87,08%), indicando boa preservação da fala nesses cenários. Para ruídos pontuais, como bebê e cachorro, os índices caem para cerca de 53% a 59%. Isso sugere que o algoritmo do Krisp também é mais eficaz contra ruídos estáveis. O nível médio da fala foi preservado com pouca atenuação.

**4.2.3 Teams**

**Quadro 3** – Métricas do Teams

| **Ruído** | **Similaridade Espectral com Original (%)** | **Nível Médio (Base) [dBFS]** | **Nível Médio (Filtrado) [dBFS]** |
| --- | --- | --- | --- |
| bebe | 37,55% | -17,37 | -19,66 |
| cachorro | 42,99% | -17,01 | -18,5 |
| caminhão | 70,96% | -17,8 | -18,86 |
| multidão | 67,79% | -18,31 | -18,77 |
| obras | 46,15% | -18,04 | -19,27 |

**Fonte**: Os autores.

O Microsoft Teams apresentou desempenho moderado na preservação da fala. Houve boa similaridade espectral apenas com ruídos como caminhão (70,96%) e multidão (67,79%), mas resultados baixos com ruídos pontuais como bebê (37,55%) e obras (46,15%). A atenuação do nível médio de áudio também foi mais acentuada em comparação às outras plataformas. Isso sugere um filtro mais agressivo, que pode comprometer partes da fala.

**4.2.4 Telegram**

**Quadro 4** – Métricas do Telegram

| **Ruído** | **Similaridade Espectral com Original (%)** | **Nível Médio (Base) [dBFS]** | **Nível Médio (Filtrado) [dBFS]** |
| --- | --- | --- | --- |
| bebe | 46,68% | -17,37 | -22,27 |
| cachorro | 55,31% | -17,01 | -20,24 |
| caminhão | 58,29% | -17,8 | -22,03 |
| multidão | 68,26% | -18,31 | -21,83 |
| obras | 69,7% | -18,04 | -21,54 |

**Fonte**: Os autores.

O Telegram demonstrou desempenho moderado na supressão de ruídos, com destaque para ruídos contínuos como obras e multidão (acima de 68% de similaridade espectral). Entretanto, houve queda significativa no nível médio de áudio, indicando uma filtragem agressiva que pode reduzir a inteligibilidade da fala, especialmente em ruídos mais intensos.

**4.2.5 Whatsapp**

**Quadro 5** – Métricas do Whatsapp

| **Ruído** | **Similaridade Espectral com Original (%)** | **Nível Médio (Base) [dBFS]** | **Nível Médio (Filtrado) [dBFS]** |
| --- | --- | --- | --- |
| bebe | 66,1% | -17,37 | -21,61 |
| cachorro | 66,75% | -17,01 | -21,99 |
| caminhão | 60,55% | -17,8 | -22,59 |
| multidão | 60,49% | -18,31 | -23,11 |
| obras | 62,33% | -18,04 | -21,85 |

**Fonte**: Os autores.

O WhatsApp apresentou resultados mais equilibrados, com boa similaridade espectral para ruídos como cachorro (66,75%) e bebê (66,1%), mantendo consistência em diferentes cenários. No entanto, assim como no Telegram, houve forte atenuação do nível médio do sinal, o que pode impactar a clareza do áudio transmitido.

**4.2.6 Zoom**

**Quadro 6** – Métricas do Zoom

| **Ruído** | **Similaridade Espectral com Original (%)** | **Nível Médio (Base) [dBFS]** | **Nível Médio (Filtrado) [dBFS]** |
| --- | --- | --- | --- |
| bebe | 95,91% | -17,37 | -17,11 |
| cachorro | 96,11% | -17,01 | -17,38 |
| caminhão | 96,98% | -17,8 | -17,45 |
| multidão | 98,39% | -18,31 | -18,1 |
| obras | 96,63% | -18,04 | -17,33 |

**Fonte**: Os autores.

O Zoom apresentou desempenho excelente em todos os cenários, com similaridade espectral acima de 95% para todos os tipos de ruído, indicando que o conteúdo da fala foi preservado com alta fidelidade. Além disso, os níveis médios dos áudios filtrados permaneceram próximos aos da base, o que sugere uma supressão de ruído eficiente sem comprometer a inteligibilidade do sinal. Esses resultados destacam o Zoom como uma das plataformas mais equilibradas entre preservação de conteúdo e redução de ruído.

# CONCLUSÃO

O presente estudo avaliou a eficácia dos algoritmos de supressão de ruído de cinco plataformas de comunicação por voz, por meio de testes com diferentes tipos de ruídos e análise visual e quantitativa dos sinais. A metodologia permitiu compreender como cada tecnologia lida com interferências sonoras comuns, como choros de bebê, latidos e ruídos urbanos, com foco na preservação da fala.

Os resultados evidenciaram diferenças significativas entre os algoritmos, que se mostraram mais eficazes diante de ruídos contínuos. Enquanto plataformas como Zoom e Discord com Krisp apresentaram melhor desempenho geral, outras adotaram filtros mais agressivos, reduzindo ruídos, mas com impacto na inteligibilidade da voz.

Conclui-se que a avaliação técnica de algoritmos de supressão de ruído é essencial para orientar escolhas informadas, principalmente em contextos que exigem clareza na comunicação. Os achados deste trabalho contribuem para o entendimento das limitações e capacidades dessas tecnologias, além de incentivarem avanços futuros na área.

# REFERÊNCIAS BIBLIOGRÁFICAS

GATTI, Pedro Henrique de Medeiros. Estudo comparativo entre duas abordagens para remoção de ruído em sinais de fala: Filtro de Wiener e Autoencoders. (Trabalho de Conclusão – Ciências da Computação). - Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas (Ibilce), São José do Rio Preto, 2024.

TSUJI JUNIOR, Nelson; SILVA, Fernando Soares da; MENDONÇA, Jane Corrêa Alves; CINTRA, Renato Fabiano. A videoconferência na realização de um concurso docente durante a pandemia de Covid-19. Desafio Online, v. 11, n. 2, p. 334–352, maio/ago. 2023. Disponível em: https://drive.google.com/file/d/1nSj01paFIsgh2-T0pkmVzcTWVBUkFnBY/view?usp=sharing. Acesso em: 13 abr. 2025.

SALLES, Laís Monique Mendes. Desenvolvimento de um dispositivo de medição de ruído com base no sistema OpenWrt. 2017. Dissertação (Mestrado em Engenharia Mecânica – Especialidade: Mecânica dos Sólidos) – Faculdade de Engenharia de Ilha Solteira, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Ilha Solteira, 2017. Disponível em: http://hdl.handle.net/11449/154132. Acesso em: 20 jun. 2025.

EMBARCADOS. Introdução ao Processamento Digital de Sinais (DSP) – Parte 1. Disponível em: <https://embarcados.com.br/introducao-ao-processamento-digital-de-sinais-dsp-parte-1/>. Acesso em: 13 abr. 2025.

EMBARCADOS. Processamento Digital de Sinais (DSP) – Parte 2. Disponível em: <https://embarcados.com.br/processamento-digital-de-sinais-dsp-parte-2/>. Acesso em: 13 abr. 2025.

ESTATÍSTICA FÁCIL. O que é filtro de Wiener? Aplicação e teoria. Disponível em: <https://estatisticafacil.org/glossario/o-que-e-filtro-de-wiener-aplicacao-e-teoria/>. Acesso em: 13 abr. 2025.

TELECO. Tutorial: Antenas Inteligentes – LMS. Disponível em: <https://www.teleco.com.br/tutoriais/tutorialsmartant/pagina_5.asp>. Acesso em: 13 abr. 2025.

MCPHEE, Brian; RAFFEL, Colin; LIANG, Dawen; ELLIS, Daniel P. W.; MC­VICAR, Matt; BATTENBERG, Eric; NIETOK, Oriol. librosa: audio and music signal analysis in Python. *Proceedings of the 14th Python in Science Conference (SciPy 2015)*, 2015.

ROHDE & SCHWARZ USA, Inc. *Meaning of 0 dBFS and 0 dBr*. FAQ, 2025. Disponível em: https://www.rohde-schwarz.com. Acesso em: 21 jun. 2025.

Statista, 2024a Number of monthly active Facebook users worldwide as of 1st quarter 2024. Disponível em: <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/>. Acesso em: 12 abr. 2025.

Statista, 2024b Most popular global mobile messenger apps as of April 2024, based on number of monthly active users. Disponível em: <https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/>. Acesso em: 12 abr. 2025.

Statista, 2024c Number of monthly active Instagram users from January 2013 to April 2024. Disponível em: <https://www.statista.com/statistics/253577/number-of-monthly-active-instagram-users/>. Acesso em: 12 abr. 2025.

KHOZHAYEV, Zulfidin. A Practical Guide to Spectrogram Analysis for Audio Signal Processing. ArXiv, 14 mar. 2024.