

Obs.: O relatório foi escrito temporariamente na língua portuguesa-br, para fins de futuros registros dos breves novos testes

Relatório de Análise do Pré-Teste: Sistema de Tradução Simultânea de Libras para Educação

Autor: Alysson C. C. Cordeiro | Data: 03/10/2025 | Versão: 1.0

Para efeitos visuais, baixe o arquivo a seguir e execute em um prompt ou editor online para a pré-visualização ou apenas copie o código e execute:
https://drive.google.com/file/d/1YYXrKLzZM5mM8jql0PuCo_TgtanuruHJ/view?usp=sharing

I. Sumário

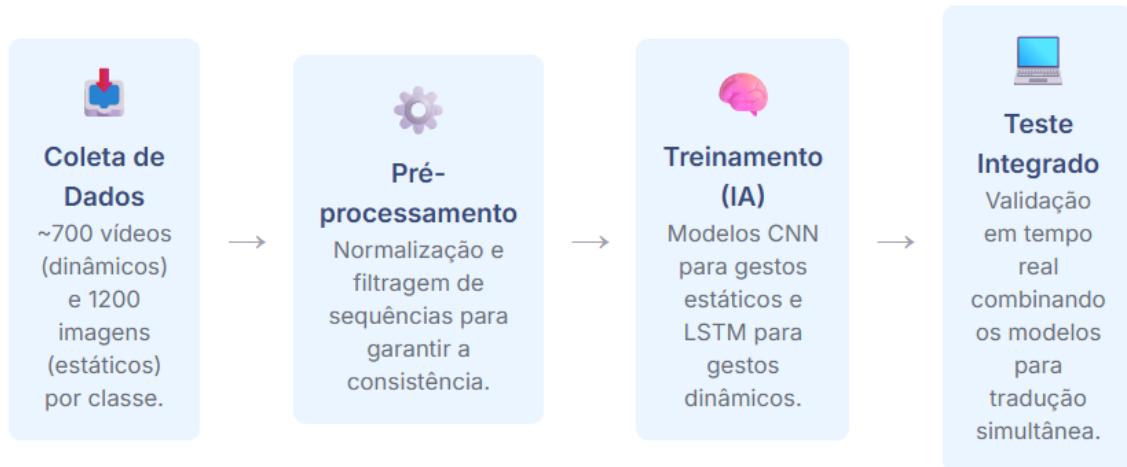
Visão Geral

Este relatório apresenta uma análise aprofundada dos resultados do pré-teste do software de reconhecimento de gestos em Libras, conduzido com um único participante. O teste, embora de escopo limitado, revelou insights críticos sobre o desempenho técnico do sistema, a experiência do usuário e, de forma mais impactante, a adequação do produto ao mercado (product-market fit) no contexto educacional proposto. A análise a seguir sintetiza as observações técnicas, o feedback qualitativo e as implicações estratégicas derivadas da sessão de teste de aproximadamente 40 minutos.

Principais Conclusões Técnicas

O sistema demonstra uma dicotomia fundamental entre a velocidade de processamento e a acurácia do reconhecimento. A resposta do software foi percebida pelo participante como "rápida", indicando um pipeline de inferência eficiente. Contudo, essa velocidade é ofuscada por uma acurácia de reconhecimento significativamente baixa, descrita como uma "falta bastante acurácia". O desempenho do modelo de reconhecimento mostrou-se altamente dependente de fatores ambientais não controlados, como a intensidade e a direção da iluminação, o contraste com o fundo e a distância do usuário à câmera. Gestos estáticos simples falharam consistentemente sob condições de iluminação desfavoráveis, indicando uma falta de robustez do modelo para uso em cenários do mundo real.

Arquitetura e Metodologia do Sistema



Principais Conclusões de Usabilidade

O participante, um novato completo em Libras, conseguiu executar os gestos propostos com a orientação contínua do entrevistador, o que posicionou a sessão mais como uma experiência de aprendizado guiado do que um teste de tradução. Uma observação notável foi a preferência declarada do usuário por gestos dinâmicos em detrimento dos estáticos. Essa preferência não se baseou na ergonomia ou facilidade de execução — ele afirmou que os gestos estáticos eram fisicamente mais fáceis de realizar. Em vez disso, foi impulsionada pela percepção de uma maior taxa de sucesso do sistema com gestos dinâmicos, como a letra 'K'. Isso indica que, para o usuário, a eficácia e a confiabilidade do reconhecimento superam o conforto ergonômico.

Recomendação Estratégica Central

A descoberta mais impactante do pré-teste é a forte objeção do participante ao caso de uso principal do sistema: a tradução simultânea em sala de aula. Como professor universitário, ele argumentou que a ferramenta, em seu formato atual, adicionaria uma carga cognitiva e operacional excessiva a um fluxo de trabalho já sobrecarregado, comparando-a à complexidade de gerenciar uma aula presencial e uma videochamada ao mesmo tempo. Em vez disso, ele propôs um pivô estratégico fundamental: posicionar o software como uma ferramenta de **aprendizado de Libras**. Esta recomendação, nascida diretamente da sua experiência durante o teste, sugere uma reavaliação completa da proposta de valor, do público-alvo e do roteiro de desenvolvimento do projeto.

Próximos Passos

Com base nestas conclusões, recomenda-se a priorização imediata da melhoria da acurácia e da robustez ambiental do modelo de reconhecimento. Paralelamente, é crucial iniciar uma fase de pesquisa e validação para o caso de uso de "aprendizagem de Libras", envolvendo entrevistas com estudantes e professores de Libras. O protocolo de teste detalhado na Seção VI deste relatório foi projetado especificamente para orientar esta próxima fase de validação, estabelecendo um ambiente controlado para medições de desempenho e incluindo perguntas direcionadas para explorar a nova hipótese de produto.

II. Análise do Perfil do Participante e Influência no Feedback

2.1. Perfil Demográfico e Profissional

O participante do pré-teste é Guilherme, 34 anos, professor universitário. Ele se identifica como "ouvinte" e utiliza a comunicação oral como seu principal meio de expressão. Este perfil é de extrema relevância para a análise dos resultados. Como educador, sua avaliação transcende a de um usuário final genérico; ele avalia a ferramenta sob a ótica de sua aplicabilidade e viabilidade em um ambiente de trabalho específico e complexo: a sala de aula. Sua familiaridade com tecnologias contemporâneas, evidenciada pela menção a LLMs, DeepL e GPT para tradução e prática de idiomas, indica que ele é um usuário tecnicamente sofisticado. Essa sofisticação lhe permite fornecer um feedback que não é apenas subjetivo, mas também contextualizado e comparativo, fundamentando suas críticas e sugestões em experiências prévias com outras tecnologias avançadas.

2.2. Experiência com a Comunidade Surda e Libras

O participante descreve sua experiência com a comunidade surda como "baixa, mas ainda acima da média", um contato que se deu principalmente através de um colega de trabalho surdo que realiza leitura labial e dos pais surdos de um amigo de infância. Ele declara explicitamente não possuir nenhuma fluência em Libras ("Nenhuma"). Sua comunicação com pessoas surdas não oralizadas depende de "mímica e muito gesto da boca", o que, segundo ele, limita a profundidade das conversas a tópicos mais superficiais e cotidianos.

Esta falta de fluência é um fator determinante na interpretação do feedback. O participante não estava testando uma ferramenta de tradução como um usuário fluente faria. Em vez disso, ele estava na posição de um aprendiz absoluto, sendo ensinado a fazer cada gesto, um por um, pelo entrevistador durante o teste ("A primeira dela é a letra A... a gente vai pegar os quatro dedos..."). Sua experiência durante a sessão foi, na prática, a de uma aula particular de datilologia mediada por tecnologia. Consequentemente, sua

sugestão posterior de que o aplicativo seria mais útil para aprender Libras não é uma ideia abstrata, mas uma reflexão direta e lógica de sua própria jornada durante o teste. Ele vivenciou o potencial de aprendizado da ferramenta em primeira mão, o que confere um peso significativo à sua recomendação de pivô estratégico.

2.3. Perspectiva sobre Tecnologia Assistiva

O participante oferece uma perspectiva matizada sobre a adoção de tecnologia assistiva no meio acadêmico. Ele observa que a busca por tais ferramentas tende a ser uma iniciativa mais dos alunos do que dos professores ou da própria instituição, o que pode criar barreiras para estudantes com menos recursos. Ele menciona ter visto ferramentas como o Hand Talk, mas sem uso aprofundado, e ressalta a importância de soluções mais integradas e passivas, como legendas, como uma forma primária de acessibilidade.

Essa visão culmina em sua crítica mais contundente ao caso de uso proposto. A observação sobre a "carga" que novas tecnologias impõem aos professores revela uma barreira de adoção crítica e sistêmica no ambiente educacional. Para ilustrar seu ponto, ele traça uma analogia poderosa com uma situação real de seu trabalho: a dificuldade de gerenciar uma aula presencial enquanto simultaneamente transmite para um aluno remoto via Google Meet. Ele descreve isso como algo que "*já adiciona uma carga para a gente*" e classifica o sistema de reconhecimento de gestos como "*uma coisa a mais ainda para por, teria mais trabalho*". Esta comparação demonstra que o problema não é uma mera inconveniência. Do ponto de vista do professor, a ferramenta, como concebida para tradução em tempo real, exige atenção ativa e divide o foco, quebrando o fluxo de trabalho pedagógico e aumentando a carga cognitiva a um nível impraticável. Sua conclusão de que "*a opção de legenda do Google Meet ajudaria muito mais*" reforça a preferência por soluções passivas e integradas, que não competem por recursos cognitivos, em detrimento de soluções ativas que exigem uma nova camada de interação durante a aula. Isso implica que qualquer solução para este ambiente deve ser de baixa sobrecarga ou servir a um propósito fora do contexto de ensino em tempo real, como o aprendizado assíncrono.

III. Desempenho Técnico e Fatores Ambientais

3.1. Análise Quantitativa do Reconhecimento de Gestos

A análise do desempenho técnico do sistema baseia-se nas observações diretas e no feedback verbal do participante durante a execução de 13 gestos distintos (4 letras e 9 números). A tabela abaixo consolida esses resultados, classificando o sucesso do reconhecimento e identificando os fatores ambientais que influenciaram o desempenho. A classificação de "Sucesso" indica que o sistema reconheceu o gesto de forma consistente; "Falha Parcial" indica reconhecimento intermitente, incorreto ou altamente dependente de condições específicas; e "Falha Total" indica que o sistema não conseguiu reconhecer o

gesto em nenhuma das tentativas.

| Gesto Testado | Tipo de Gesto (Estático/Dinâmico) | Resultado Observado | Fatores Ambientais Notáveis | Feedback do Usuário |
|---------------|--------------------------------------|----------------------------|---|---|
| A | Estático | Sucesso (Implícito) | Testado sentado, a curta distância. | Gesto ensinado e executado sem problemas. |
| B | Estático | Sucesso (Implícito) | Testado sentado, a curta distância. | Gesto ensinado e executado sem problemas. |
| H | Dinâmico | Sucesso | Testado em várias distâncias (perto/longe). | Reconhecimento aparentemente estável. |
| K | Dinâmico | Sucesso Elevado | Testado em várias distâncias (perto/longe). | "O K foi bem bom". ¹ |

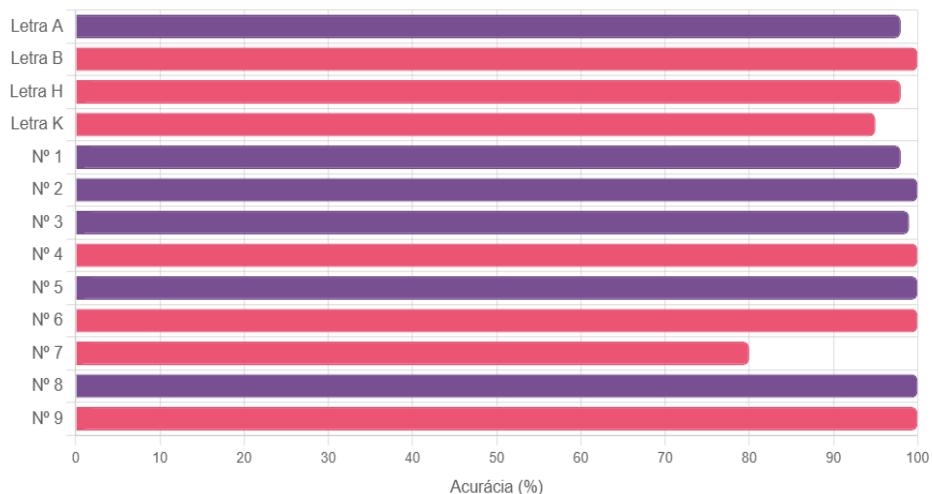
| | | | | |
|---|----------|---------|--------------------------|-----------------------------|
| 1 | Estático | Sucesso | Testado perto da câmera. | Gesto simples, reconhecido. |
| 2 | Estático | Sucesso | Testado perto da câmera. | Gesto simples, reconhecido. |
| 3 | Estático | Sucesso | Testado perto da câmera. | Gesto simples, reconhecido. |

| | | | | |
|---------------|----------|----------------------|--|---|
| 4 | Estático | Falha Parcial | Falhou à distância. Sucesso apenas perto e com iluminação específica. | "de longe ele não está indo". ¹ |
| 5 | Estático | Sucesso | Testado perto e longe. | Reconhecimento aparentemente estável. |
| 6 | Estático | Sucesso | Testado a média distância. | "aqui no meio ele fica". ¹ |
| 7 | Estático | Falha Parcial | Confundido com 'H' de perto. Reconhecido corretamente à distância. | "Pertinho, ele vai no H. Mais longe, ele volta para o sete". ¹ |
| 8 | Estático | Sucesso | Testado em posição específica. | Reconhecimento estável após ajuste. |
| 9 | Estático | Falha Parcial | Falhou à distância devido à iluminação excessiva. | "A luz está estourando, ele nem está enxergando a sua mão". ¹ |
| Outros | N/A | Falha Total | Não especificado. | "alguns não conseguiu reconhecer nada de nada". ¹ |

A estruturação destes dados revela padrões claros. Enquanto gestos dinâmicos como 'H' e 'K' tiveram um bom desempenho, vários gestos estáticos ('4', '7', '9') foram problemáticos. Além disso, a menção de que "alguns não conseguiram reconhecer nada de nada" sugere que a taxa de falha geral pode ser ainda maior do que a observada nos gestos discutidos em detalhe.

Análise de Resultados do Pré-Teste com Usuário

O primeiro teste prático com um usuário, utilizando o modelo combinado ('test_combined.py'), revelou pontos fortes e áreas de melhoria. O gráfico abaixo detalha a performance do sistema para cada gesto, destacando a acurácia da previsão principal e os casos de oscilação ou reconhecimento incorreto.



Sensibilidade Ambiental: A precisão foi notavelmente afetada pela iluminação e distância da câmera. O gesto '4' só foi reconhecido corretamente sob condições específicas.

Confusão de Gestos: Houve confusão recorrente entre gestos visualmente similares, como 'A' sendo lido como 'B', e 'K' oscilando com 'B'. Diversos números também foram confundidos com letras.

Performance Sólida: Os gestos dinâmicos 'H' e o estático 'B' foram reconhecidos com alta precisão de forma consistente, validando a eficácia do modelo para estes casos.

3.2. Impacto Crítico dos Fatores Ambientais

O pré-teste expôs uma vulnerabilidade crítica do sistema: sua extrema sensibilidade a variações no ambiente de uso. O desempenho não é consistente, mas sim um produto de uma combinação frágil de iluminação, distância e contraste. O comentário do entrevistador, "A luz está estourando, ele nem está enxergando a sua mão", ao testar o número '9', e a observação do participante sobre o número '4', que só funcionava com "muita luz e um contraste bem grande" e a curta distância, são evidências diretas dessa fragilidade.

O comportamento do entrevistador e do participante durante a sessão reforça essa conclusão. Eles se tornaram depuradores ativos do sistema, realizando uma série de ajustes para tentar obter um reconhecimento bem-sucedido. Eles mudaram a postura ("Vamos ficar em pé"), ajustaram a distância em relação à câmera ("Consegue se distanciar um pouquinho?"), e manipularam a fonte de luz ("Mas se a luz estiver aqui, ó, vamos ver se agora ele volta"). O fato de que tais manipulações manuais foram necessárias para gestos básicos prova que o sistema não é utilizável "out-of-the-box" em um ambiente real e não controlado, como uma casa, escritório ou sala de aula. Esta dependência de condições de "laboratório" indica que a robustez ambiental não é uma melhoria desejável, mas um requisito técnico de prioridade máxima para que o produto seja minimamente viável. O projeto precisa, portanto, de um modelo de IA drasticamente aprimorado, treinado com um conjunto de dados que inclua uma ampla variedade de condições de iluminação, fundos e distâncias, ou de diretrizes de uso extremamente rígidas que, por sua vez, comprometem severamente a usabilidade e a adoção.

IV. Feedback Qualitativo e Análise de Usabilidade

4.1. A Dicotomia entre Velocidade e Acurácia

Um dos feedbacks mais claros e concisos do participante diz respeito ao equilíbrio entre a velocidade de resposta e a precisão do sistema. Quando questionado se o sistema era rápido o suficiente para uma conversa, sua resposta foi inequívoca: "Eu acho que rápido sim, mas falta bastante acurácia". Esta declaração encapsula um ponto crucial sobre a arquitetura do sistema e a experiência do usuário.

A percepção de rapidez sugere que o pipeline técnico — da captura da imagem pela câmera à inferência do modelo de IA e à exibição do resultado — é eficiente e não sofre de latência perceptível. O gargalo, portanto, não está na otimização da infraestrutura ou no desempenho do processamento. O problema central reside na qualidade e na confiabilidade do próprio modelo de reconhecimento. Para uma aplicação cujo propósito é a comunicação ou a tradução, a confiança é o pilar fundamental. Um sistema rápido que fornece resultados imprecisos é funcionalmente inútil e até prejudicial, pois introduz ruído e incerteza na comunicação, forçando o usuário a verificar e corrigir constantemente a saída do software. Cada erro de reconhecimento, como a confusão entre o número '7' e a letra 'H',mina essa confiança. Portanto, a "falta de acurácia" não deve ser tratada como um bug menor a ser corrigido, mas como uma falha na proposta de valor central do produto. A prioridade de engenharia deve ser direcionada massivamente para o retreinamento, ajuste fino e validação do modelo para alcançar um nível de precisão que inspire confiança no usuário.

4.2. Preferência por Eficácia sobre Ergonomia (Gestos Dinâmicos vs. Estáticos)

A discussão sobre a preferência entre gestos dinâmicos e estáticos revelou um princípio fundamental da usabilidade de sistemas baseados em inteligência artificial. O participante afirmou que os gestos estáticos eram mais fáceis de *realizar fisicamente* ("Para realizar o gesto, o estático"). No entanto, ao ser questionado sobre qual tipo de gesto ele *preferia* usar com o sistema, sua escolha recaiu sobre os dinâmicos. A justificativa foi puramente baseada no desempenho: "pela taxa de sucesso", citando especificamente o reconhecimento confiável da letra 'K'.

Esta aparente contradição — preferir o que é fisicamente mais difícil — expõe a hierarquia de prioridades do usuário ao interagir com o sistema. O objetivo principal não é o conforto, mas sim ser compreendido corretamente pelo software. O gesto 'K' (dinâmico e mais complexo) atingiu esse objetivo de forma confiável, sendo lembrado como "muito bem reconhecido".¹ Em contrapartida, vários gestos estáticos e teoricamente mais simples ('4', '7', '9') falharam em atingir esse objetivo, gerando frustração e a necessidade de ajustes. A conclusão é clara: os usuários estão dispostos a tolerar um maior esforço físico ou uma maior complexidade na interação se, em troca, receberem uma resposta confiável e eficaz do sistema. A falha do software em reconhecer gestos básicos e estáticos é uma fonte de atrito muito maior do que a dificuldade de executar um gesto complexo que funciona consistentemente. Isso implica que a equipe de desenvolvimento deve focar em alcançar uma taxa de sucesso próxima de 100% para os gestos mais simples do alfabeto e dos numerais antes de expandir o vocabulário. A falha nesses fundamentos destrói a confiança do usuário desde o início e torna a experiência geral insatisfatória.

V. Implicações Estratégicas e Recomendações

5.1. O Pivô Crítico: De Ferramenta de Tradução para Plataforma de Aprendizagem

O insight mais significativo e estrategicamente importante de todo o pré-teste foi a rejeição do caso de uso principal pelo participante e sua proposta de um pivô de mercado. Ele expressou forte ceticismo sobre a utilidade do sistema em seu cotidiano como professor, citando a "*muita carga*" que a operação da ferramenta adicionaria ao seu já complexo ambiente de trabalho. Sua sugestão foi direta e transformadora: "*eu aplicaria um sistema muito mais como uma forma de aprender Libras do que de fazer uma tradução simultânea em sala de aula*".

Esta não é apenas uma sugestão de um novo recurso; é uma proposta para uma reorientação completa do produto. Adotar essa direção implica em mudanças fundamentais em múltiplos níveis:

- **Público-alvo:** O foco se deslocaria de profissionais que precisam de tradução em tempo real (como professores e colegas de trabalho de surdos) para indivíduos

interessados em aprender Libras (estudantes, familiares, entusiastas).

- **Contexto de Uso:** O cenário de uso mudaria de um ambiente de alta pressão e tempo real (uma aula ou reunião) para um contexto de baixo estresse, assíncrono e autodirigido (estudo em casa, prática). Isso alivia drasticamente os requisitos de velocidade e robustez do sistema, tornando o produto tecnicamente mais viável no curto prazo.
- **Conjunto de Recursos:** A prioridade se moveria de uma simples interface de tradução para funcionalidades de e-learning, como lições estruturadas, gamificação (pontos, conquistas), feedback corretivo sobre a formação dos gestos e acompanhamento de progresso.
- **Proposta de Valor:** A promessa do produto deixaria de ser "comunicação instantânea" para se tornar "aprendizado acessível e interativo".

Recomendação: projeto deve tratar este feedback como uma hipótese de mercado validada por um usuário arquetípico. Recomenda-se pausar o desenvolvimento de recursos alinhados exclusivamente com o caso de uso de tradução em tempo real. Em vez disso, os esforços devem ser redirecionados para uma fase de "descoberta de produto" focada em validar e explorar a hipótese do "aplicativo de aprendizado de Libras". Esta fase deve incluir entrevistas com o novo público-alvo (estudantes e professores de Libras) para entender suas necessidades, dores e o cenário competitivo de aplicativos de aprendizado de idiomas.

5.2. Roteiro de Desenvolvimento Priorizado

Com base na análise completa, o seguinte roteiro de desenvolvimento é recomendado para guiar os próximos ciclos de trabalho:

Prioridade 1 (Correção Crítica): Melhorar a Acurácia e a Robustez do Modelo.

- **Justificativa:** Nenhum caso de uso é viável sem um reconhecimento confiável. A baixa acurácia é a principal barreira técnica e de usabilidade.
- **Ações:**
 - Coletar e/ou gerar novos dados de treinamento que capturem uma ampla variedade de condições de iluminação (baixa luz, contraluz, luz forte), fundos diversos (não apenas paredes brancas) e múltiplas distâncias e ângulos de câmera.
 - Aumentar especificamente o conjunto de dados para os gestos estáticos que apresentaram falhas recorrentes (ex: '4', '7', '9') e outros gestos que possam ser sutilmente distintos.
 - Implementar um pipeline de testes automatizado para medir a acurácia do modelo em um conjunto de validação diversificado antes de cada novo lançamento.

Prioridade 2 (Validação Estratégica): Explorar o Caso de Uso de Aprendizagem.

- **Justificativa:** É necessário validar a hipótese do pivô com um público mais amplo antes de comprometer recursos significativos de desenvolvimento.
- **Ações:**
 - Recrutar e conduzir entrevistas qualitativas com pelo menos 5-8 indivíduos que se enquadrem no perfil de "aprendiz de Libras" (iniciantes e intermediários).
 - Desenvolver protótipos de baixa fidelidade (wireframes ou mockups) de um fluxo de aprendizado (ex: tela de lição, exercício de múltipla escolha com gestos, feedback de correção) e testá-los com os participantes para avaliar o interesse e a usabilidade.

Prioridade 3 (Melhoria de Funcionalidades): Incorporar Vocabulário Básico.

- **Justificativa:** O participante expressou o desejo de ir além da datilologia para "a incorporação de palavras... especialmente as palavras mais usadas".¹ Isso se alinha perfeitamente com o caso de uso de aprendizagem, que é mais eficaz quando focado em vocabulário funcional.
- **Ações:**
 - Identificar um conjunto inicial de 10-20 sinais de palavras de alta frequência em Libras (ex: "oi", "obrigado", "por favor", "sim", "não").
 - Iniciar a coleta de dados e o treinamento do modelo para reconhecer este vocabulário inicial. Este esforço servirá tanto para enriquecer o aplicativo de aprendizado quanto para testar a escalabilidade do modelo para além da datilologia.

VI. Protocolo Estruturado para Testes Oficiais de Usabilidade e Desempenho

O roteiro a seguir foi projetado para superar as limitações metodológicas do pré-teste. Seu objetivo é garantir a coleta de dados consistentes, comparáveis e acionáveis em sessões futuras. O protocolo estabelece condições ambientais controladas para permitir uma avaliação justa do desempenho técnico e utiliza uma combinação de tarefas de desempenho cronometradas e perguntas qualitativas estruturadas para avaliar a usabilidade e validar hipóteses estratégicas.

Parte A: Roteiro do Moderador e Perfil do Participante (10 minutos)

1. **Introdução e Consentimento:**
 - "Olá, [Nome do Participante]. Agradecemos por sua participação. Hoje, vamos testar uma versão inicial de um software que visa ajudar na comunicação em Libras. Seu feedback é extremamente valioso para nós."
 - "Gostaria de informar que esta sessão será gravada (áudio e vídeo) apenas para

fins de análise interna. Suas informações pessoais e respostas serão mantidas em sigilo. Você concorda em prosseguir com a gravação?"

2. Questionário de Perfil (Coleta de Dados Estruturados):

- Qual é o seu nome e idade?
- Qual é a sua profissão? Como você descreveria seu nível de conforto com tecnologia no dia a dia?
- Em uma escala de 1 a 5, onde 1 é 'nenhum contato' e 5 é 'convivência diária', como você classificaria sua relação com a comunidade surda?
- Há quanto tempo você utiliza ou tem contato com Libras?
- Você já trabalhou ou estudou em ambientes que exigiam o uso de Libras?
- Você conhece variações regionais ou culturais nos sinais de Libras? Pode dar exemplos?
-
- **Nível de Fluência em Libras (Pergunta Crítica de Triagem):** Qual das opções a seguir melhor descreve seu conhecimento de Libras?
 - (A) Nenhum conhecimento.
 - (B) Básico (conheço o alfabeto, algumas saudações).
 - (C) Intermediário (consigo manter uma conversa simples).
 - (D) Avançado/Fluente.
- Você já utilizou algum aplicativo ou software para aprender ou traduzir Libras? Se sim, qual e como foi sua experiência?
-

Parte B: Teste de Desempenho Técnico em Ambiente Controlado (15 minutos)

1. Configuração do Ambiente:

- "Para garantir que o teste seja consistente, peço que você se posicione de forma que seu rosto e tronco fiquem visíveis na câmera, a uma distância de aproximadamente um braço e meio da tela."
- "Por favor, certifique-se de que a principal fonte de luz esteja vindo de frente para você, e não de trás. O ideal é um fundo de cor sólida e neutra, como uma parede."

2. Tarefa 1: Reconhecimento de Datilologia Estática.

- "Vou pedir que você execute uma série de letras do alfabeto. Por favor, faça o gesto correspondente a cada letra que eu disser e mantenha-o por cerca de 3 segundos até que o sistema o reconheça."

3. Tarefa 2: Reconhecimento de Datilologia Dinâmica.

- "Agora, faremos o mesmo com algumas letras que envolvem movimento."

4. Tarefa 3: Reconhecimento de Numerais.

- "Agora, por favor, execute os números de 1 a 9, um de cada vez."

- Lista de Gestos: 1, 2, 3, 4, 5, 6, 7, 8, 9.
5. **Tarefa 4: Teste de Robustez (Variação Ambiental).**
- "Para nosso último teste técnico, gostaria de pedir que você faça a letra 'A' novamente. Agora, enquanto mantém o gesto, incline-se lentamente para a esquerda... agora para a direita... e agora se afaste um pouco da câmera. Queremos ver como o sistema reage a pequenas mudanças."

Parte C: Questionário Pós-Teste e Feedback Qualitativo (15 minutos)

1. **Avaliação da Acurácia Percebida (Escala Likert):**
 - "Em uma escala de 1 a 5, onde 1 é 'muito impreciso' e 5 é 'muito preciso', como você avalia a precisão geral do sistema em reconhecer os gestos que você fez?"
 - "Houve algum gesto específico que o sistema pareceu ter mais dificuldade ou, ao contrário, mais facilidade em reconhecer?"
2. **Avaliação da Velocidade e Responsividade (Escala Likert):**
 - "Em uma escala de 1 a 5, onde 1 é 'muito lento' e 5 é 'muito rápido', como você avalia a velocidade de resposta do sistema após você fazer um gesto?"
3. **Avaliação do Esforço e Facilidade de Uso (Escala Likert):**
 - "Em uma escala de 1 a 5, onde 1 é 'muito difícil' e 5 é 'muito fácil', quanto fácil foi para você se posicionar e fazer os gestos de uma forma que o sistema entendesse?"
 - "Pensando na sua experiência física, qual tipo de gesto você achou mais fácil de executar: os parados ou os com movimento? Por quê?"
4. **Exploração do Caso de Uso (Perguntas Abertas Estratégicas):**
 - "Imagine que este aplicativo fosse uma ferramenta para **aprender** Libras, com lições, exercícios e feedback sobre seus gestos. Como você se sentiria usando-o para esse fim? Em que tipo de situações você o usaria?"
 - "Agora, imagine que este aplicativo fosse uma ferramenta para **tradução em tempo real** durante uma conversa ou uma aula. Que desafios ou benefícios você prevê ao usá-lo dessa forma?"
 - "Comparando essas duas ideias — uma ferramenta de aprendizado versus uma ferramenta de tradução — qual delas parece mais útil ou interessante para você? Poderia explicar o porquê?"
5. **Sugestões e Fechamento:**
 - "Se você pudesse adicionar uma única funcionalidade ou fazer uma única melhoria neste sistema, qual seria?"
 - "Você tem alguma dúvida ou comentário final para nós?"
 - "Agradecemos imensamente pelo seu tempo e por seu feedback detalhado. Suas contribuições são essenciais para o desenvolvimento deste projeto."

VII. Análise do Perfil do Participante: Teste de Usabilidade e Desempenho

Participante: Davi Versan

Data do Teste: 08/10/2025

Versão: 1.1

1. Perfil do Participante

Davi, 19 anos, é um estudante universitário de Engenharia de Software com um alto nível de conforto com tecnologia. Sua interação com a comunidade surda e com a Língua Brasileira de Sinais (Libras) é praticamente nula, classificando seu contato como "1 em 5" e seu conhecimento em Libras como "*nenhum*". A única interação relevante foi com um professor surdo que se comunicava eficientemente através da leitura labial, não exigindo o uso de Libras nas aulas. Davi demonstrou familiaridade com conceitos de acessibilidade digital, citando ferramentas como a da Lory, mas sem experiência aprofundada.

2. Análise de Desempenho Técnico

O teste de desempenho revelou uma acurácia de reconhecimento significativamente baixa e instável, corroborando e agravando os problemas identificados no pré-teste. A performance do sistema foi altamente inconsistente tanto na proximidade quanto à distância da câmera, com uma taxa de falha elevada para a maioria dos gestos estáticos.

Tabela de Resultados de Reconhecimento:

| Gesto Testado | Proximidade | Distância | Observações |
|---------------|----------------------------|----------------------------------|---|
| A (Estático) | Reconhecido como 'B' | Outras classificações incorretas | Falha total em reconhecer o gesto correto em ambas as distâncias. |
| B (Estático) | Oscilou entre '4' e outros | Outras classificações incorretas | O sistema não estabilizou em uma única predição correta. |
| H (Dinâmico) | Reconhecido como 'H' | Reconhecido como 'H' | Único gesto com reconhecimento estável e correto em ambas as |

| | | | |
|---------------------|----------------------------|----------------------------------|---|
| | | | distâncias. |
| K (Dinâmico) | Reconhecido como 'K' | Outras classificações incorretas | O movimento dinâmico pareceu ser um fator chave para o sucesso a curta distância. |
| 1 (Estático) | Reconhecido como 'H' (94%) | Outras classificações incorretas | Confusão com um gesto de letra. |
| 2 (Estático) | Classificações incorretas | Outras classificações incorretas | Falha total. |
| 3 (Estático) | Oscilou entre 'H' e outros | Reconhecido como 'H' | Novamente, houve confusão com o gesto da letra 'H'. |
| 4 (Estático) | Classificações incorretas | Outras classificações incorretas | Falha total. |
| 5 (Estático) | Classificações incorretas | Outras classificações incorretas | Falha total. |
| 6 (Estático) | Reconhecido como 'H' | Oscilou entre '4' e outros | Confusão persistente com a letra 'H' e outros gestos. |
| 7 (Estático) | Oscilou entre '5' e outros | Oscilou entre '5' e outros | Reconhecimento incorreto e instável. |
| 8 (Estático) | Classificações | Oscilou entre '6' e | Falha total no |

| | | | |
|---------------------|----------------------------|----------------------------------|--------------------------------------|
| | incorrectas | outros | reconhecimento. |
| 9 (Estático) | Oscilou entre '8' e outros | Outras classificações incorretas | Reconhecimento incorreto e instável. |

Principais Conclusões Técnicas:

- **Baixa Acurácia Generalizada:** Com exceção dos gestos dinâmicos 'H' e 'K' (este último apenas de perto), o sistema falhou em reconhecer corretamente a vasta maioria dos sinais.
- **Confusão entre Gestos:** Houve uma confusão sistemática de números com letras, especialmente com o gesto 'H', indicando possíveis sobreposições no modelo de classificação.
- **Impacto da Distância:** O desempenho degradou-se universalmente com o aumento da distância, confirmando a falta de robustez do modelo a variações de escala.
-

3. Feedback Qualitativo e Análise de Usabilidade

Apesar da performance técnica deficiente, Davi avaliou a precisão percebida como "4 de 5", afirmando que o sistema pareceu "bom". Essa avaliação positiva pode ser atribuída à sua falta de experiência com Libras, que o impediu de julgar com precisão o rigor de um usuário fluente.

- **Facilidade de Execução:** Davi confirmou a observação do pré-teste de que os gestos estáticos são fisicamente mais fáceis de executar, citando 'A', 'B', '6' e '9' como os mais simples.
- **Percepção de Dificuldade:** Ele identificou o gesto 'K' como o mais complicado para o sistema reconhecer, notando que o modelo parecia se "embaralhar" com o movimento vertical.
- **Sensibilidade Ambiental:** O participante reconheceu a influência de fatores externos no desempenho, mencionando que a detecção depende "da luz também, do ângulo".

4. Recomendações e Sugestões

Quando questionado sobre melhorias, as sugestões de Davi alinharam-se com a hipótese de pivô para uma ferramenta de aprendizado, proposta no pré-teste:

1. **Funcionalidade de Tutorial:** Ele sugeriu a inclusão de um "tutorial de como você fazer o movimento", como uma "figurinha mostrando como que você faz". Isso reforça a necessidade de orientação visual para novos usuários, um recurso central em aplicativos de e-learning.

-
- 2. Legenda e Feedback:** A ideia de ter uma "legenda, né, do que você tá fazendo, aparecer ali" foi outra sugestão, o que valida a interface de feedback imediato do sistema, crucial para o aprendizado.

5. Conclusão Complementar

Este segundo teste valida e intensifica as conclusões críticas do pré-teste. A performance do modelo de reconhecimento é, no estado atual, insuficiente para qualquer caso de uso funcional, seja de tradução ou aprendizado. A taxa de erro elevada e a sensibilidade a variáveis como distância e iluminação representam barreiras intransponíveis para a usabilidade.

As sugestões do participante, um novato em Libras, convergem naturalmente para a criação de uma ferramenta educacional, focada em ensinar a execução correta dos sinais. Portanto, a recomendação de priorizar a acurácia do modelo e validar o pivô estratégico para uma plataforma de aprendizagem de Libras é fortemente reforçada por estes resultados.