

Automação de bill of materials na indústria têxtil utilizando inteligência artificial e processamento multimodal

Automation of bill of materials in the textile industry using artificial intelligence and multimodal processing

Jose Francisco de Souza Fonseca¹
Matheus Felipe Duarte Dionísio²
Matheus Mathias Rocha Lucio de Moraes³
Rafael da Silva Cosme⁴
Rafael Martins da Silva⁵
Robinson Luis de Souza Alves⁶
Ivanilson França Vieira Júnior⁷
Wagner de Oliveira⁸

Resumo

Este trabalho propõe uma solução tecnológica para otimizar o fluxo produtivo na indústria de confecção, visando agilizar o preenchimento do Bill of Materials (BOL) na etapa de prototipagem. O estudo aborda a substituição do registro manual de operações de costura por um sistema automatizado de captura de dados audiovisuais. O objetivo central é desenvolver uma aplicação para registro multimodal das tarefas, integrando ferramentas de visão computacional e processamento de linguagem natural. A metodologia emprega a linguagem Python e Inteligência Artificial, utilizando o modelo Whisper para Reconhecimento Automático de Fala (ASR) e busca semântica para a classificação e estruturação dos comandos operacionais. Os resultados demonstram que o sistema viabiliza o preenchimento automático do BOL e a criação de um acervo audiovisual técnico, promovendo a

¹ Discente do Curso de Formação Inicial e Continuada (Curso FIC) em Residência Tecnológica em Software Embarcado, na modalidade a distância. e-mail: franciscojosefrancisco341@gmail.com

² Discente do Curso de Formação Inicial e Continuada (Curso FIC) em Residência Tecnológica em Software Embarcado, na modalidade a distância. e-mail: matheus.felipe.du@gmail.com

³ Discente do Curso de Formação Inicial e Continuada (Curso FIC) em Residência Tecnológica em Software Embarcado, na modalidade a distância. e-mail: mathias.matheus76@gmail.com

⁴ Discente do Curso de Formação Inicial e Continuada (Curso FIC) em Residência Tecnológica em Software Embarcado, na modalidade a distância. e-mail: rafaelcosme282@gmail.com

⁵ Discente do Curso de Formação Inicial e Continuada (Curso FIC) em Residência Tecnológica em Software Embarcado, na modalidade a distância. e-mail: rafaeldevrn@gmail.com.br

⁶ Docente do Curso de Formação Inicial e Continuada (Curso FIC) em Residência Tecnológica em Software Embarcado, na modalidade a distância. e-mail: robinson.alves@ifrn.edu.br

⁷ Docente do Curso de Formação Inicial e Continuada (Curso FIC) em Residência Tecnológica em Software Embarcado, na modalidade a distância. e-mail: ivanilson.junior@ifrn.edu.br

⁸ Docente do Curso de Formação Inicial e Continuada (Curso FIC) em Residência Tecnológica em Software Embarcado, na modalidade a distância. e-mail: wagner.oliveira@ifrn.edu.br

padronização dos processos e a redução do tempo de revisão. Conclui-se que a automação baseada no reconhecimento de voz e vídeo constitui uma alternativa eficaz aos métodos tradicionais, aumentando a confiabilidade da informação técnica sem depender exclusivamente de processos manuais.

Palavras-chave: Automação Têxtil; Inteligência Artificial; Reconhecimento de Fala; Sistemas Embarcados; Processamento Multimodal.

Abstract

This paper proposes a technological solution to optimize the production flow in the garment industry, aiming to streamline the filling of the Bill of Materials (BOL) during the prototyping stage. The study addresses the replacement of manual recording of sewing operations with an automated system for audiovisual data capture. The main objective is to develop an application for multimodal recording of tasks, integrating computer vision and natural language processing tools. The methodology employs Python language and Artificial Intelligence, using the Whisper model for Automatic Speech Recognition (ASR) and semantic search for the classification and structuring of operational commands. The results demonstrate that the system enables the automatic filling of the BOL and the creation of a technical audiovisual collection, promoting process standardization and reducing revision time. It is concluded that automation based on voice and video recognition constitutes an effective alternative to traditional methods, increasing the reliability of technical information without relying exclusively on manual processes.

Keywords: Textile Automation; Artificial Intelligence; Speech Recognition; Embedded Systems; Multimodal Processing.

1 INTRODUÇÃO

A indústria têxtil contemporânea vivencia um processo de transformação digital, buscando integrar tecnologias de automação e análise de dados para otimizar fluxos produtivos e garantir a rastreabilidade da manufatura. Neste cenário, a etapa de prototipagem de peças de vestuário destaca-se como um ponto crítico, pois envolve a definição de processos e insumos que nortearão a produção em larga escala. Um dos elementos centrais desta etapa é o Bill of Materials (BOL), documento técnico que reúne todas as operações, tipos de tecido e componentes necessários para a confecção da roupa.

Apesar da evolução tecnológica no setor, o preenchimento e a manutenção do BOL em muitas indústrias ainda ocorrem de maneira predominantemente manual. Conforme observado no contexto de uma grande empresa do setor têxtil parceira desta pesquisa, peças complexas sofrem constantes alterações durante a prototipagem. Esse dinamismo exige que os analistas revisem e ajustem manualmente a documentação repetidas vezes, um processo suscetível a erros humanos, que consome tempo significativo da equipe de engenharia. Além disso, a

complexidade de certas operações de costura gera dúvidas recorrentes no chão de fábrica, evidenciando a carência de registros visuais que padronizam a execução das tarefas.

Para mitigar tais desafios, a literatura técnica aponta para o uso de interfaces multimodais e Inteligência Artificial. Lopes Filho (2022) destaca que modelos modernos baseados em redes neurais profundas ampliaram significativamente a precisão do Reconhecimento Automático de Fala (Automatic Speech Recognition - ASR), tornando viável seu uso em ambientes industriais. Simultaneamente, o avanço dos Grandes Modelos de Linguagem (Large Language Models - LLMs) permitiu que sistemas computacionais transcendam a simples correspondência de palavras-chave, sendo capazes de compreender nuances linguísticas e realizar inferências complexas em linguagem natural, conforme aponta a Stryker (2025).

A presente pesquisa justifica-se pela necessidade de preencher a lacuna existente entre a execução física da costura e o registro digital da operação. A automatização deste processo promete não apenas reduzir o tempo despendido com o preenchimento manual e a revisão do BOL, como também criar um acervo de conhecimento audiovisual, facilitando treinamentos futuros e garantindo a padronização da qualidade. A relevância prática do estudo reside na aplicação direta de tecnologias emergentes para resolver um gargalo real de eficiência operacional, enquanto sua relevância teórica consiste na exploração de arquiteturas que combinam visão computacional e processamento de linguagem natural em ambientes fabris.

Diante do exposto, o objetivo geral consiste em desenvolver uma solução automatizada de software que realize o preenchimento do BOL e registre em vídeo as operações de costura. Os objetivos específicos incluem capturar áudio e vídeo das operações em tempo real, interpretar comandos verbais utilizando modelos de IA para identificar as operações executadas.

2 FUNDAMENTAÇÃO TEÓRICA

Esta fundamentação teórica apresenta os conceitos e tecnologias que embasam o desenvolvimento do trabalho, abordando a computação embarcada e o hardware de captura, o processamento de áudio e vídeo, e os métodos de reconhecimento automático de fala aplicados a ambientes industriais.

2.1. Computação Embarcada e Hardware de Captura

A computação embarcada refere-se à integração de sistemas computacionais diretamente em dispositivos físicos para a execução de tarefas dedicadas em tempo real. No

contexto desta pesquisa, destaca-se o uso do Single Board Computer (SBC) Raspberry Pi 5. Segundo a Newark (2024), esta plataforma representa uma evolução significativa em capacidade de processamento e conectividade, sendo apta para aplicações que demandam captura de vídeo de alta definição e comunicação com APIs externas sem depender de infraestrutura de servidores locais robustos.

2.2. Processamento de Áudio e Vídeo

Para a manipulação eficiente dos fluxos de mídia, utilizou-se o framework FFmpeg. Segundo Wilbert (2024), esta ferramenta é referência mundial em codificação e transmissão de vídeo, permitindo converter formatos de arquivo e enviar transmissões em tempo real via protocolos como RTMP. Devido a essa flexibilidade e robustez, a ferramenta torna-se ideal para operar em ambientes industriais, garantindo a integridade das informações capturadas mesmo em dispositivos com recursos limitados de hardware.

2.3. Reconhecimento Automático de Fala

O Reconhecimento Automático de Fala (Automatic Speech Recognition - ASR) é o processo de conversão de linguagem falada em texto processável. Modelos contemporâneos baseados em redes neurais profundas (Deep Neural Networks) têm superado as limitações históricas de precisão. Lopes Filho (2022) evidencia que arquiteturas baseadas em Transformers (como o Wav2Vec2) superam o desempenho de redes recorrentes, apresentando maior adaptabilidade para sistemas comerciais. Contudo, a robustez contra ruídos intensos ainda representa um desafio em certas implementações. Visando superar essa barreira no contexto de chão de fábrica, este trabalho adota o modelo Whisper, da OpenAI, que demonstra alta resiliência na transcrição multilíngue, lidando eficazmente com ruídos de fundo típicos de ambientes de confecção (OPENAI, 2022).

2.4. Embeddings Semânticos e Recuperação de Informação

Para interpretar os comandos transcritos, aplica-se o conceito de embeddings, que são representações numéricas de palavras ou frases em um espaço vetorial multidimensional. Conforme definido pela Cloudflare ([s.d.]), essa técnica permite capturar relações semânticas e contextuais entre termos, possibilitando que o sistema identifique a similaridade entre o que foi dito pelo operador e as operações técnicas cadastradas no sistema, superando a rigidez da busca por palavras-chave exatas.

2.5. Modelos de Linguagem de Grande Escala (LLMs)

A etapa final de processamento utiliza Modelos de Linguagem de Grande Escala (*Large Language Models* - LLMs). Segundo Stryker (2025), tais sistemas, treinados com vastos volumes de dados, possuem a capacidade de realizar inferências complexas e compreender nuances linguísticas. Na arquitetura proposta, a LLM atua como um agente de validação contextual em uma abordagem RAG (*Retrieval-Augmented Generation*), refinando os resultados da busca vetorial para garantir que a operação identificada seja tecnicamente coerente com a sequência de montagem da peça.

3 METODOLOGIA

A presente pesquisa caracteriza-se como aplicada e exploratória, envolvendo o desenvolvimento de um protótipo funcional de hardware e software para validação em ambiente industrial real. A abordagem metodológica foi estruturada em três etapas principais: definição da arquitetura do sistema, implementação do ambiente de hardware e software, e procedimentos de teste e validação de campo.

3.1. Arquitetura do Sistema

O fluxo operacional proposto opera em uma arquitetura cliente-servidor. Na borda (*edge*), uma estação de trabalho embarcada realiza a interface com o usuário e a captura de mídia. O processo inicia-se com a autenticação do operador via leitura de código de barras, seguida pela seleção do projeto de costura na interface *touchscreen*. Após a confirmação, o sistema ativa a captura simultânea de vídeo e áudio.

Como exibido na Figura 1, os dados capturados são transmitidos para um servidor central, onde ocorre o processamento assíncrono em um *pipeline* de Inteligência Artificial composto por quatro estágios: (1) Transcrição do áudio (ASR); (2) Vetorização do texto (*embeddings*); (3) Busca por similaridade em banco de dados vetorial; e (4) Análise contextual via *Large Language Model* (LLM) para estruturação do documento *Bill of Materials* (BOL) em formato JSON.

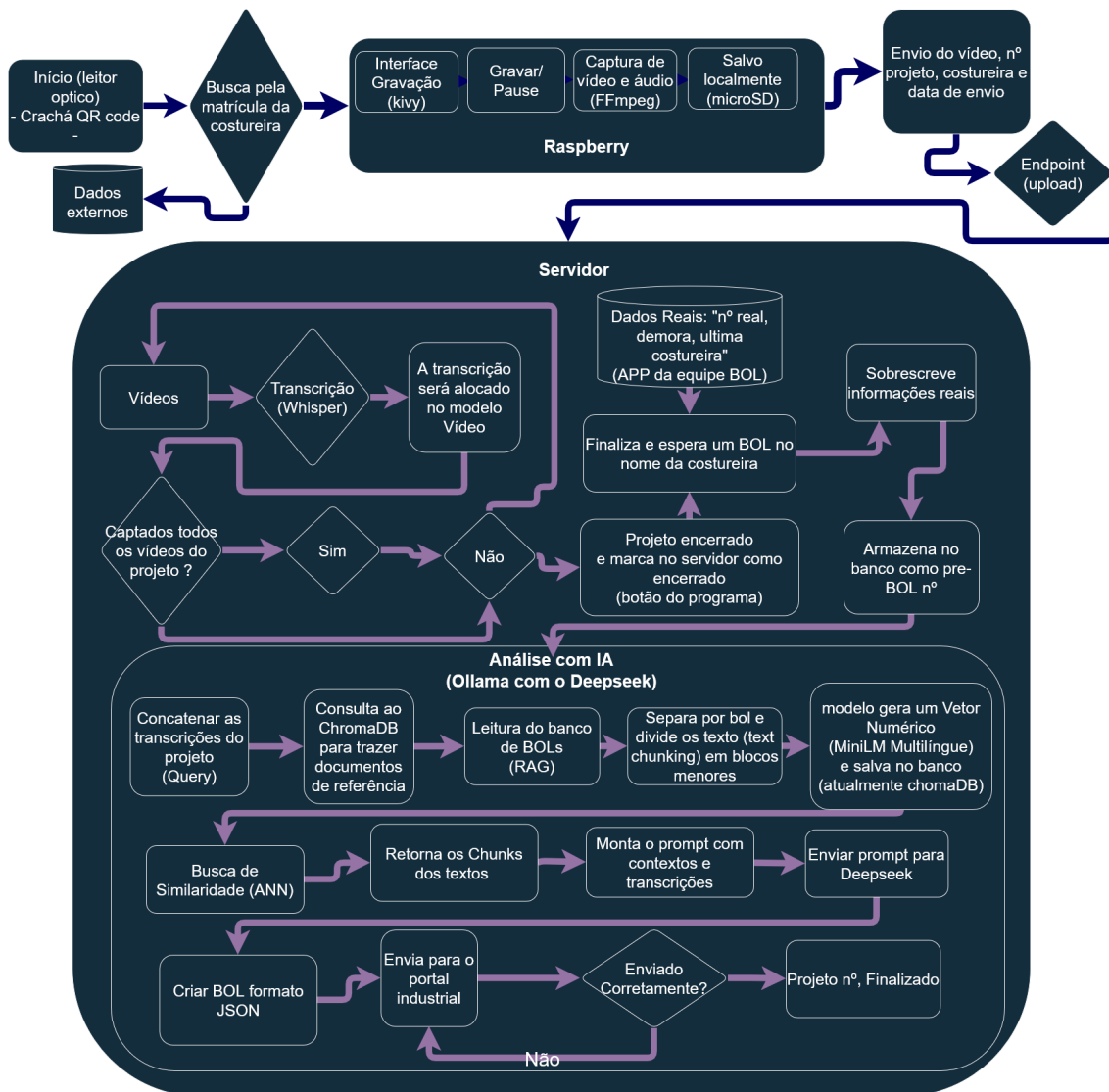


Figura 1: Diagrama de blocos do fluxo de dados e processamento.

Fonte: Autores (2026).

3.2. Materiais e Hardware Utilizados

Para a composição da estação de trabalho embarcada, foram selecionados componentes que garantem desempenho e robustez térmica. O núcleo de processamento é uma *Raspberry Pi 5*, escolhida por sua capacidade de lidar com fluxos de vídeo em alta definição. O sistema é alimentado por uma fonte proprietária de 27W para garantir estabilidade elétrica aos periféricos.

A interface homem-máquina (IHM) é realizada através de um monitor *touchscreen*, eliminando a necessidade de teclado e mouse. A captura visual utiliza câmeras USB Logitech, enquanto a captura sonora é realizada por um microfone direcional Rode. A escolha deste microfone específico foi metodologicamente crítica para isolar a voz do operador e minimizar

a interferência do ruído mecânico das máquinas de costura industriais.

3.3. Ferramentas de Software

O desenvolvimento do *software* utilizou a linguagem *Python* como base, devido ao seu vasto ecossistema de bibliotecas para automação e IA. A interface gráfica (GUI) foi construída com o *framework Kivy* (Kivy, 2026), operando em modo quiosque para restringir o acesso do operador apenas às funções do sistema. Para o gerenciamento de mídia, utilizou-se a biblioteca *FFmpeg* (FFMPEG, 2026), responsável pela captura, compressão e transmissão dos fluxos audiovisuais. O ambiente de desenvolvimento foi padronizado utilizando containers *Docker*, garantindo a reprodutibilidade da aplicação entre os ambientes de desenvolvimento e produção.

3.4. Procedimentos de Teste e Validação

O processo foi estruturado em três fases distintas e sequenciais, assegurando que cada etapa validasse pré-requisitos essenciais para a seguinte, conferindo robustez à solução final entregue.

A Fase 1 foi sobre os testes funcionais realizados com base de dados sintética, gravada pela própria equipe de pesquisa, para validar a lógica de transcrição e integração dos módulos de software.

A Fase 2 focou na avaliação da estabilidade do hardware sob carga contínua. A Raspberry Pi 5 foi submetida a ciclos de operação de 24 e 48 horas ininterruptas para monitoramento de temperatura e verificação de falhas na sincronia de vídeo (dropped frames).

A Fase 3 consistiu na implantação do protótipo na linha de produção da indústria parceira. Foram coletadas amostras reais de costureiras durante o turno de trabalho para aferir a precisão do modelo Whisper frente ao ruído ambiente real e validar a usabilidade da interface gráfica. Os dados quantitativos obtidos nesta fase final, incluindo a comparação entre as operações reais e as detectadas, estão detalhados na Seção 4.

4 RESULTADOS E DISCUSSÕES

A implementação do sistema foi concluída seguindo as fases estipuladas na metodologia, resultando em um protótipo funcional integrado ao ambiente fabril. A análise dos resultados foi segmentada em três vertentes: usabilidade da interface e captura, eficiência do processamento inteligente e estabilidade sistêmica sob estresse.

4.1. Implementação do Ambiente Embarcado e Interface

A etapa inicial de montagem física confirmou a adequação da *Raspberry Pi 5* para a tarefa. O uso de um sistema de dissipação ativa permitiu que o *hardware* operasse com estabilidade térmica, mesmo executando a captura de vídeo em alta definição simultaneamente à gestão da interface gráfica. O sistema operacional configurado em modo quiosque garantiu a segurança da aplicação, restringindo o acesso do usuário apenas às funcionalidades permitidas.

Em relação à interface com o usuário (Frontend), desenvolvida em *Kivy*, os testes de campo indicaram alta aceitação por parte dos operadores. O *design* com botões ampliados e alto contraste (modo escuro), como mostrado na Figura 3, facilitou a interação via tela sensível ao toque, minimizando erros de operação. O fluxo de autenticação, que permite o uso de crachá com código de barras, como é visto na Figura 2, agilizou o início das atividades, vinculando corretamente a autoria das operações aos metadados do vídeo.

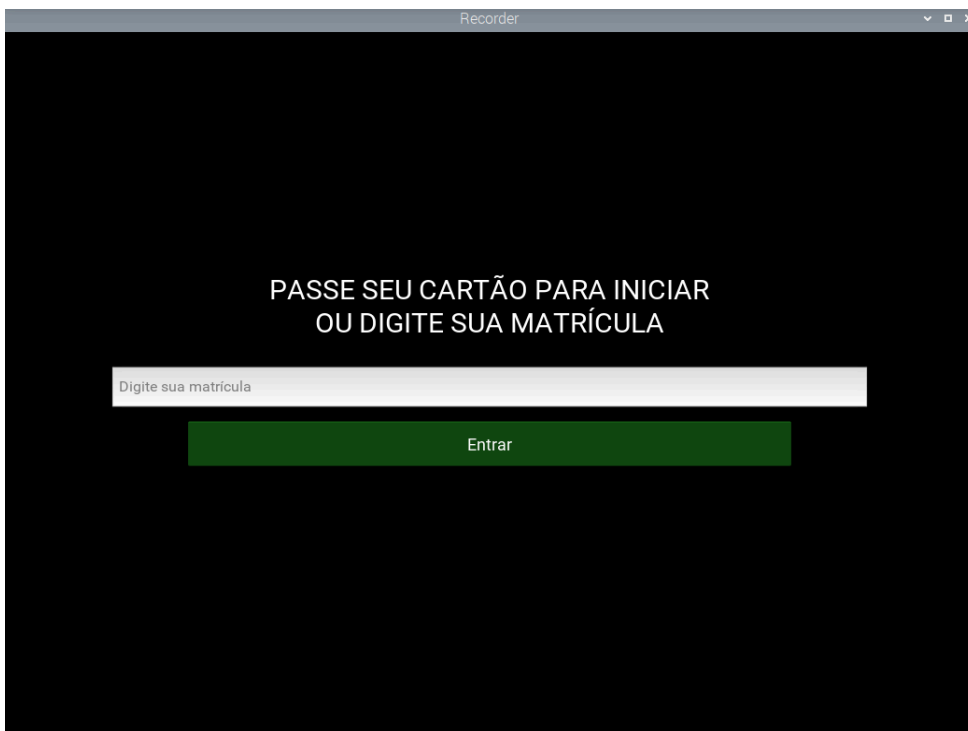


Figura 2: Tela de autenticação com suporte a leitura de código de barras.

Fonte: Autores (2026).

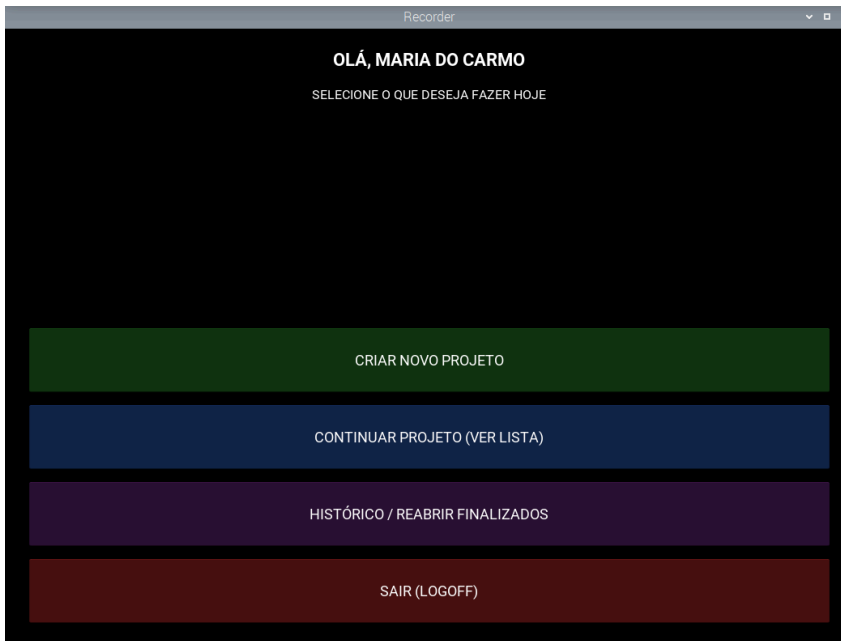


Figura 3: Painel principal de seleção de fluxo de trabalho.

Fonte: Autores (2026).

A interface de gravação fornece *feedback* visual em tempo real (Figura 4), permitindo a costureira ajustar o enquadramento da peça antes de iniciar a descrição verbal.

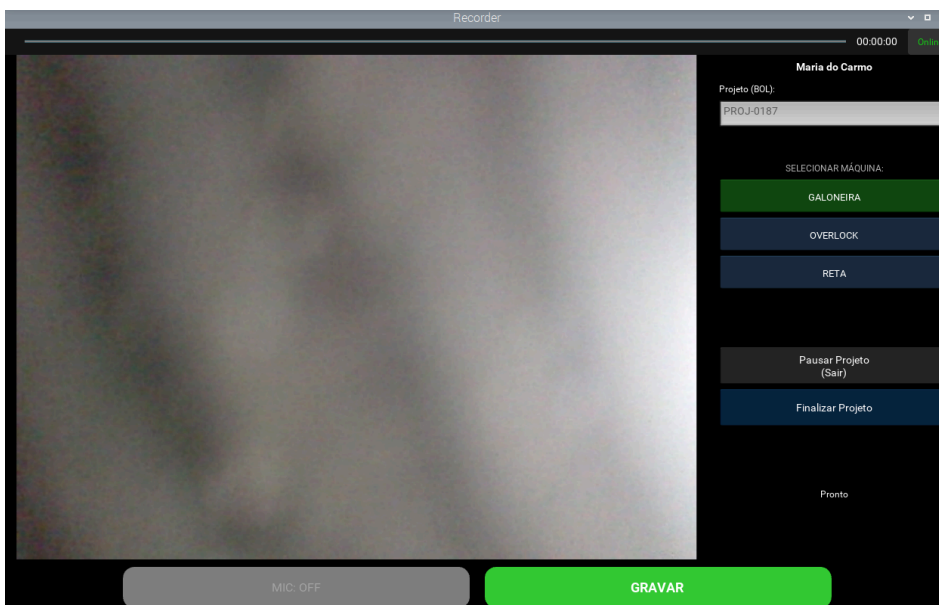


Figura 4: Interface de captura ativa com controles operacionais.

Fonte: Autores (2026).

4.2. Eficiência do Reconhecimento de Fala e Processamento

Para quantificar a eficácia do sistema em ambiente fabril (Fase 3 da metodologia), foi realizada uma comparação direta entre os comandos verbais transcritos e as operações padronizadas do Bill of Materials. A Tabela 1 apresenta os dados de validação de um ciclo completo de prototipagem, confrontando a operação real realizada pela costureira (Gabarito) com a operação classificada pelo sistema após o processamento via embeddings e LLM.

Tabela 1: Comparativo entre operações reais e detectadas pelo sistema.

Ordem	Transcrição (Input)	ID Real (Gabarito)	Operação Real (Gabarito)	ID Detectado (Sistema)	Operação Detectada (Sistema)	Confiança	Validação
0.0	Iniciando a costura de união superior agora.	4	Costura de União Superior	4	Costura de União Superior	0.79	Correto
1.0	Fazendo a aplicação de identificação térmica na peça.	3	Aplicação de Identificação Térmica	3	Aplicação de Identificação Térmica	0.91	Correto
2.0	Ok, procedendo com montagem de colarinho.	6	Montagem de Colarinho	6	Montagem de Colarinho	0.92	Correto
3.0	Vou realizar a fixação de colarinho ao corpo aqui.	12	Fixação de Colarinho ao Corpo	12	Fixação de Colarinho ao Corpo	0.82	Correto
4.0	Terminando a etapa anterior, agora é pesponto de acabamento do decote.	9	Pesponto de Acabamento Decote	5	Costura de Fechamento Lateral	0.71	Incorreto
5.0	Executando inserção de reforço interno conforme ficha.	11	Inserção de Reforço Interno	11	Inserção de Reforço Interno	0.83	Correto
6.0	Iniciando a travetamento de reforço lateral agora.	2	Travetamento de Reforço Lateral	8	Aplicação de Aviamentos	0.74	Incorreto
7.0	Fazendo a aplicação de aviamentos na	8	Aplicação de Aviamentos	8	Aplicação de Aviamentos	0.82	Correto

	peça.							
8.0	Ok, procedendo com aplicação de aviamentos.	10	Aplicação de Aviamentos	de 10	Aplicação de Aviamentos	0.74	Correto	
9.0	Vou realizar a costura de fechamento lateral aqui.	5	Costura de Fechamento Lateral	de 5	Costura de Fechamento Lateral	0.83	Correto	
10.0	Terminando a etapa anterior, agora é costura de fechamento lateral.	7	Costura de Fechamento Lateral	de 7	Costura de Fechamento Lateral	0.82	Correto	
11.0	Executando costura de fechamento lateral conforme ficha.	1	Costura de Fechamento Lateral	de 1	Costura de Fechamento Lateral	0.96	Correto	
TOTAL	-	-	-	-	-	0.828	83.3%	

Fonte: Autores (2026).

A análise dos dados na Tabela 1 demonstra que o sistema obteve uma taxa de acerto de 83,3% na amostra validada (10 acertos em 12 interações). Observa-se que os acertos apresentaram um índice de confiança médio de 0.828, enquanto os dois erros registrados ("Pesponto de Acabamento" e "Travetamento de Reforço") ocorreram em instâncias onde a confiança do modelo caiu para a faixa de 0,71 a 0,74. Essa queda correlaciona-se com a complexidade semântica das frases, sugerindo que o modelo de similaridade ainda encontra dificuldades em diferenciar operações com terminologias técnicas muito específicas quando há ruído de fundo.

Aprofundando a discussão técnica, a utilização de um microfone direcional profissional elevou significativamente a Relação Sinal-Ruído (SNR), isolando a voz do operador frente ao ruído contínuo das máquinas, o que foi determinante para a alta assertividade nas operações de costura padrão.

Contudo, o fenômeno de "alucinação" do modelo em momentos de silêncio ou baixa modulação de voz persistiu como um desafio pontual. Apesar dessas ocorrências, a integração com o Large Language Model (LLM) na etapa de pós-processamento atuou como um filtro eficaz na maioria dos casos, corrigindo inconsistências contextuais e garantindo a estruturação correta do Bill of Materials (BOL). No que tange ao desempenho temporal, a busca vetorial

(vector search) apresentou latência desprezível, confirmando a eficiência da arquitetura proposta.

4.3. Testes de Estabilidade e Discussão

Os testes de estresse foram fundamentais para determinar os limites operacionais do sistema embarcado. No cenário padrão de operação, simulando um turno de 24 horas contínuas, o sistema manteve-se estável, sem perda de dados ou superaquecimento, validando a robustez do *Single Board Computer* escolhido.

Entretanto, ao estender a simulação para 48 horas ininterruptas, detectou-se instabilidade na sincronização de áudio e vídeo, manifestada pela duplicação de quadros (*dropped frames*) no processo de gravação via *FFmpeg*.

A discussão desses resultados aponta para uma causalidade híbrida: embora o *hardware* suporte a carga de processamento, a configuração de *software* para captura contínua de longuíssima duração requer otimização no gerenciamento de *buffer*. Comparando o esperado com o alcançado, o sistema atende plenamente à demanda operacional real (turnos de trabalho diários), posicionando a falha de 48 horas como um limite técnico de "teste de estresse" que não compromete a viabilidade comercial do produto, desde que respeitados os ciclos de reinicialização ou manutenção diária.

A solução proposta, portanto, supera a barreira da simples digitalização, oferecendo uma ferramenta ativa de gestão do conhecimento. A combinação de *embeddings* semânticos com LLMs demonstrou ser capaz de mitigar a variabilidade linguística dos operadores (gírias ou termos regionais), garantindo que a documentação técnica gerada mantenha a padronização exigida pela engenharia de produto.

5 CONCLUSÃO

O desenvolvimento desta pesquisa confirma que a aplicação de sistemas embarcados integrados à Inteligência Artificial Generativa constitui uma solução viável e robusta para a automação documental na indústria têxtil. O objetivo principal de substituir o preenchimento manual do *Bill of Materials* (BOL) por um sistema de captura audiovisual inteligente é atingido, demonstrando que a arquitetura proposta opera de forma estável e contínua em ambiente de produção real.

A análise dos dados permite concluir que a integração entre *hardware* de borda e processamento em nuvem, orquestrada por uma API proprietária, garante a integridade e a segurança dos dados fabris. Uma das principais descobertas do estudo reside na constatação

de que a eficiência do reconhecimento de fala em chão de fábrica depende tanto da sofisticação dos modelos de linguagem quanto da qualidade física da captura de áudio. A arquitetura implementada valida o uso de *embeddings* semânticos como ferramenta eficaz para mitigar a variabilidade linguística dos operadores, assegurando a padronização técnica necessária para a engenharia de produtos.

Como contribuição prática, o sistema transcende a simples automação burocrática, entregando um acervo visual estruturado que serve como base para treinamento e gestão do conhecimento. No entanto, o estudo apresenta limitações relacionadas ao *hardware* de captura de imagem, cuja qualidade óptica provisória restringe a análise detalhada de costuras microscópicas. Além disso, a sensibilidade do modelo de transcrição a ruídos ambientais estáticos indica a necessidade de refinamentos na lógica de ativação da gravação.

Para trabalhos futuros, sugere-se a implementação de gatilhos manuais de áudio (*push-to-talk*) ou mecanismos de detecção de voz (*Voice Activity Detection*) via *hardware* para eliminar a captação de ruídos espúrios. Recomenda-se também a integração direta dos dados estruturados aos sistemas de gestão do ciclo de vida do produto (PLM), eliminando etapas intermediárias de validação e consolidando a transformação digital do setor.

REFERÊNCIAS

CLOUDFLARE. **O que são embeddings no aprendizado de máquina?** Cloudflare Learning, [s.d.]. Disponível em: <https://www.cloudflare.com/pt-br/learning/ai/what-are-embeddings/>. Acesso em: 13 jan. 2026.

DIDÁTICA TECH. **Tudo sobre o framework Kivy para iniciantes!** [S. l.], [s.d.]. Disponível em: <https://didatica.tech/tudo-sobre-o-framework-kivy-para-iniciantes/>. Acesso em: 15 jan. 2026.

FFMPEG. **FFmpeg: A complete, cross-platform solution to record, convert and stream audio and video.** Disponível em: <https://ffmpeg.org/>. Acesso em: 29 jan. 2026.

KIVY. **Kivy: Open source Python library for rapid development of applications.** Disponível em: <https://kivy.org/>. Acesso em: 29 jan. 2026.

LOPES FILHO, Antonio Alves. **Análise comparativa de redes profundas para reconhecimento automático de fala em português.** 2022. 51 p. Monografia (MBA em Inteligência Artificial e Big Data) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2022. Disponível em: https://bdta.abcd.usp.br/directbitstream/7716b2eb-7006-4fa2-b06c-a858bd42691d/Antonio%20Alves_TCC_ASR_COMPARISON_OFICIAL%20%287%29_207623.pdf. Acesso em: 16 jan. 2026.

NEWARK. **Raspberry Pi 5: Revisão Completa e Análise de Desempenho**. Embarcados, 20 jun. 2024. Disponível em: <https://embarcados.com.br/raspberry-pi-5-revisao-completa-e-analise-de-desempenho/>. Acesso em: 16 jan. 2026.

OPENAI. **Introducing Whisper**. OpenAI Research, 2022. Disponível em: <https://openai.com/research/whisper>. Acesso em: 16 jan. 2026.

STRYKER, Cole. **O que é LLM (grandes modelos de linguagem)?** IBM Think, 8 dez. 2025. Disponível em: <https://www.ibm.com/br-pt/think/topics/large-language-models>. Acesso em: 15 jan. 2026.

WILBERT, Max. **FFmpeg Live Streaming – What Professional Broadcasters Need to Know [2025 Update]**. Blog Dacast, 6 jun. 2025. Disponível em: <https://www.dacast.com/blog/how-to-broadcast-live-stream-using-ffmpeg/>. Acesso em: 15 jan. 2026.