

Monitorização do comportamento ergonómico de operadores na indústria 5.0 através de visão computacional baseada em Inteligência Artificial

PEDRO PAULO CAMPOS SANTOS

Dissertação para obtenção do grau de Mestre em:
Engenharia Eletrotécnica e de Computadores

Trabalho realizado sob a orientação de

Prof. Paulo Leitão

Prof.^a Flávia Pires

Prof. Claiton de Oliveira

Bragança

Novembro, 2025

Monitorização do comportamento ergonómico de operadores na indústria 5.0 através de visão computacional baseada em Inteligência Artificial

PEDRO PAULO CAMPOS SANTOS

Dissertação para obtenção do grau de Mestre em:
Engenharia Eletrotécnica e de Computadores

Trabalho realizado sob a orientação de

Prof. Paulo Leitão

Prof.^a Flávia Pires

Prof. Claiton de Oliveira

Bragança

Novembro, 2025

Dedicatória

Dedico este trabalho a minha família e amigos.

Agradecimentos

Agradeço aos meus orientadores, Prof. Paulo Leitão, Prof.^a Flávia Pires e Prof. Claiton Oliveira pelos conselhos, ideias, dedicação e paciência durante todo este trabalho.

Agradeço a minha família por sempre me apoiar, me ajudar, zelar por mim e se fazer presente mesmo com a distância. Se cheguei até aqui, é por que tenho vocês.

Aos meus amigos e amigas que compartilharam momentos comigo durante esta jornada, agradeço por cada vitória, cada café tomado juntos, pelas risadas compartilhadas, histórias contadas e simplesmente por me fazerem companhia.

Resumo

A presente dissertação propõe o desenvolvimento de um sistema não invasivo para monitorização do comportamento de operadores dentro do contexto da ergonomia, alinhado aos princípios da Indústria 5.0, que coloca o bem-estar humano no centro dos processos produtivos. O sistema visa identificar posturas inadequadas e prevenir Lesões Músculo-Esqueléticas Relacionadas com o Trabalho (LMERT), utilizando técnicas de Inteligência Artificial e Visão Computacional. Para tal, foi implementado um módulo de deteção de pose baseado no modelo MediaPipe Pose e uma câmara de profundidade Intel RealSense D415, que permite calcular ângulos posturais tridimensionais em tempo real. O software foi desenvolvido em Python, integrando módulos de filtragem temporal, persistência de eventos e uma base de dados SQLite, além de uma interface web interativa criada com Flask e WebSocket. O sistema foi validado experimentalmente por comparação com uma avaliação ergonómica conduzida por uma especialista, obtendo acurácia global de 94,6% e sensibilidade de 93,1%. Os resultados confirmam a eficácia da solução proposta na deteção automática de posturas inadequadas. O sistema contribui para o avanço da ergonomia digital e para a promoção da saúde ocupacional na Indústria 5.0, oferecendo uma ferramenta acessível, portátil e capaz de apoiar a prevenção de LMERT de forma contínua e em tempo real.

Palavras-chave: Indústria 5.0, Ergonomia, Inteligência Artificial, Visão Computacional.

Abstract

This dissertation proposes the development of a non-invasive system for monitoring operator behavior within the context of ergonomics, aligned with the principles of Industry 5.0, which places human well-being at the center of production processes. The system aims to identify inadequate postures and prevent Work-Related Musculoskeletal Disorders (WMSDs), using Artificial Intelligence and Computer Vision techniques. To this end, a pose detection module based on the MediaPipe Pose model and an Intel RealSense D415 depth camera were implemented, which allows for the calculation of three-dimensional postural angles in real-time. The software was developed in Python, integrating modules for temporal filtering, event persistence, and an SQLite database, in addition to an interactive web interface created with Flask and WebSocket. The system was experimentally validated by comparison with an ergonomic assessment conducted by a specialist, achieving an overall accuracy of 94.6% and a sensitivity of 93.1%. The results confirm the effectiveness of the proposed solution in the automatic detection of inadequate postures. The system contributes to the advancement of digital ergonomics and the promotion of occupational health in Industry 5.0, offering an accessible, portable tool capable of supporting WMSD prevention continuously and in real-time.

Keywords: Industry 5.0; Ergonomics; Artificial Intelligence; Computer Vision.

Conteúdo

1	Introdução	1
1.1	Contextualização	1
1.2	Objetivos	4
1.3	Estrutura do Documento	4
2	Estado da Arte	7
2.1	Indústria 5.0	7
2.2	Ergonomia	8
2.2.1	Distúrbios Músculo-Esqueléticos	9
2.2.2	Ferramentas de Avaliação Ergonómica	9
2.3	Visão Computacional e os Modelos de Estimaco de Pose Humana	11
2.4	Anlise Ergonmica com Viso Computacional	13
2.5	Resumo	15
3	Arquitetura e Implementaco do Sistema	17
3.1	Sistema de Anlise Ergonmica	17
3.1.1	Hardware	19
3.1.2	Software	20
3.2	Detetor de Postura	22
3.2.1	Processamento de Imagem	24
3.2.2	Projecco 2D para 3D	26
3.2.3	Clculo dos Ângulos Posturais	29

3.2.4	Filtros de Suavização	34
3.2.5	Detecções Auxiliares	37
3.2.6	Definição de Limiar e Eventos	41
3.2.7	Estrutura Geral de Dados do Sistema	42
3.3	Interface Gráfica	45
3.4	Resumo	50
4	Validação e Discussão	53
4.1	Validação do Sistema	53
4.1.1	Desempenho Geral do Sistema e Métricas Diagnósticas	53
4.1.2	Análise Detalhada da Classificação e Falsos Negativos	55
4.1.3	Valor Diagnóstico da Informação de Profundidade	57
4.1.4	Síntese dos Resultados	57
4.2	Validação de Desempenho	57
4.3	Discussão dos Resultados Obtidos	58
4.4	Considerações Éticas e Legais sobre o Uso de Imagens e Dados Pessoais	59
4.5	Resumo	60
5	Conclusão	61
5.1	Considerações Finais	61
5.2	Trabalhos Futuros	62
A	Código Fonte	71
B	Validação com Ergonomista	72

Lista de Tabelas

2.1	Níveis de ação e pontuação do <i>score</i> RULA. Adaptado de [21].	10
2.2	Critério de classificação das ferramentas ergonómicas. Adaptado de [24]. . .	11
2.3	Estudos sobre sistemas de análise ergonómica.	13
3.1	Representação dos marcadores corporais utilizados pelo <i>MediaPipe</i>	27
3.2	Limiar de ativação (<i>On</i>) e desativação (<i>Off</i>) utilizados para cada medida. .	42
3.3	Filtros aplicados e tempo de persistência adotado para cada métrica do sistema.	43
3.4	Campos armazenados na <i>baseline</i> do usuário durante o processo de calibração.	43
4.1	Matriz de confusão para deteção de posturas inadequadas ($n = 37$)	54
4.2	Métricas de desempenho do sistema de deteção ergonómica	55

Lista de Figuras

1.1	Pirâmide das Necessidades Humanas na Indústria. Adaptado de [2]	2
3.1	Diagrama do Sistema com Fluxo de Dados.	18
3.2	Ângulos monitorizados pelo sistema: (a) Inclinação frontal da cabeça; (b) Rotação lateral do tronco; (c) Inclinação frontal do tronco; (d) Inclinação lateral do tronco.	19
3.3	Diagrama de Sequência UML.	21
3.4	Representação Postura do Trabalhador e Captação da Imagem.	22
3.5	Imagem de Profundidade (a) e Imagem com Cores (b).	23
3.6	Representação Estrutural do Processamento de Imagem.	24
3.7	33 Marcadores Corporais Detetados pelo <i>MediaPipe</i> . Fonte: [36]	24
3.8	Ilustração Funcionamento Pinhole com Representação da Câmara e Objeto no Mundo Real	28
3.9	Captação dos Eixos em Relação ao Corpo do Usuário.	30
3.10	Comparação entre as Derivadas Temporais do Ângulo de <i>Pitch</i> Antes e Após a Filtragem.	35
3.11	Comparação entre as Derivadas Temporais do Ângulo de <i>Yaw</i> Antes e Após a Filtragem.	36
3.12	(a) Quadris Detetados Sobre a Mesa. (b) Quadris Corrigidos Para Fora da Mesa.	39
3.13	Diagrama Entidade–Relacionamento (DER) da Estrutura Geral do Sistema de Ergonomia.	45

3.14	Página Inicial da Interface Gráfica do Sistema.	46
3.15	Tela de Criação e Calibração de Novo Usuário.	47
3.16	Modo de Análise em Tempo Real — Visão Operador.	47
3.17	Modo de Análise em Tempo Real — Visão Gestor.	48
3.18	Tela de Gerenciamento e Exclusão de Usuários.	48
3.19	Tela de Visualização das Últimas Sessões Parte Superior.	49
3.20	Tela de Visualização das Últimas Sessões Parte Inferior.	49
3.21	Diagrama de Comunicação da Camada <i>Web Server</i>	50
4.1	Caso 1 de Falso Negativo - Leve Flexão Cervical Detetado Pela Especialista. Sistema Retornou Boa Postura	56
4.2	Caso 2 de Falso Negativo - Rotação de Cervical Detetado Pela Especialista. Sistema Retornou Boa Postura	56

Siglas

BCS Sistema de Coordenadas do Corpo. 29

CeDRI Centro de Investigação em Digitalização e Robótica Inteligente. 22

CPS Cyber-physical System. 3

CPU *Central Processing Unit*. 23, 58, 60, 61

DME Distúrbio Músculo-Esquelético. 9, 14

DORT Distúrbios Osteomusculares Relacionados ao Trabalho. 2

DT *Digital Twin*. 8

EMA *Exponential Moving Average*. 34, 51

ESTiG Escola Superior de Tecnologia e Gestão. 22

EU-OSHA *European Agency for Safety and Health at Work*. 2, 9

FPS *frames* por segundo. 20, 24, 57, 58

GDPR General Data Protection Regulation — Regulamento Geral sobre a Proteção de Dados. 59

GPU *Graphics Processing Unit*. 23, 58, 60

HPE *Human Pose Estimation*. 12–14, 22, 24, 27, 37, 41

IA Inteligência Artificial. 3, 4, 7, 8, 14, 15

IoT Internet of Things. 3, 8

IPA *Intelligent Personal Assistant*. 12, 22

IPB Instituto Politécnico de Bragança. 22

LER Lesão por Esforço Repetitivo. 2

LGPD Lei Geral de Proteção de Dados Pessoais — Lei nº 13.709, de 14 de agosto de 2018. 59

LME Lesão Músculo-Esquelética. 2

LMERT Lesão Músculo-Esquelética Relacionada com o Trabalho. 2, 3, 9, 15, 58, 61–63

OWAS *Ovaco Working Posture Analysis*. 10, 11

REBA *Rapid Entire Body Assessment*. 10, 11, 14

RULA *Rapid Upper Limb Assessment*. 10, 11, 13–15, 17, 18, 41, 53

VC Visão Computacional. 3, 4, 7, 11, 12, 23

YOLO *You Only Look Once*. 12

Capítulo 1

Introdução

Este capítulo busca contextualizar esta dissertação dentro do cenário tecnológico atual, apresentando conceitos e temas relacionados com a presente pesquisa. Também define o problema a ser tratado, o objetivo do trabalho e sua contribuição.

1.1 Contextualização

A indústria 5.0, promovida pela Comissão Europeia, é um novo paradigma industrial que coloca o bem-estar do ser humano, a resiliência e a sustentabilidade no centro do processo de produção. Esta revolução surge após o reconhecimento das limitações da indústria 4.0 evidenciadas durante a pandemia do Covid-19, justificando a necessidade de complementar a quarta revolução com um formato mais sustentável e resiliente às rupturas das cadeias de abastecimento globais [1].

O conceito de humano-centrismo (*human-centricity*) traz uma série de novas preocupações da indústria para com o trabalhador a fim de garantir o bem-estar humano. Lu et al. realizaram um estudo buscando compreender quais são as necessidades básicas do humano dentro da indústria [2]. Utilizando a Hierarquia das Necessidades de Maslow [3] como base, os autores criaram uma pirâmide de necessidade voltada à indústria em 5 níveis, como mostra a Figura 1.1 [2].

A base dessa pirâmide assenta em duas necessidades fundamentais: a segurança física

(nível 1), que protege o trabalhador de interações perigosas com máquinas, evitando acidentes de trabalho imediatos e garante seus direitos legais. E a saúde do trabalhador (nível 2), que visa garantir um ambiente de trabalho que não comprometa o bem-estar do corpo a longo prazo, evitando Lesões Músculo-Esqueléticas (LMEs). É neste segundo nível que a presente investigação se foca.



Figura 1.1: Pirâmide das Necessidades Humanas na Indústria. Adaptado de [2]

A LME é um conjunto de doenças inflamatórias do sistema locomotor que pode ser causada por diversos fatores, como tabagismo, falta de exercício físico ou fatores agravados pelo trabalho, como os movimentos repetitivos, a sobrecarga e/ou postura adotada durante a jornada de trabalho [4], [5]. Quando o principal fator de risco tem origem profissional, é designada por Lesão Músculo-Esquelética Relacionada com o Trabalho (LMERT).

Segundo a Agência Europeia para a Segurança e Saúde no Trabalho (*European Agency for Safety and Health at Work* (EU-OSHA)), a LMERT corresponde a cerca de 60% das doenças profissionais na Europa, estima-se que a cada cinco profissionais europeus, três já tenham passado por alguma LMERT ao longo da vida [6].

No Brasil, este tipo de lesão é denominado por Lesão por Esforço Repetitivo (LER) e Distúrbios Osteomusculares Relacionados ao Trabalho (DORT), que afastou em 2023 mais de 10 mil trabalhadores dos seus postos de trabalho [7]. Os setores mais afetados são

aqueles que envolvem movimentos repetitivos, posturas inadequadas, ou um esforço físico contínuo, afetando maioritariamente setores como, a indústria, principalmente as linhas de montagem e produção, setor de serviços de tele-marketing, profissionais de escritório, e o comércio, com operadores de caixa ou estoque. Isto pode levar a um grave problema de saúde pública e ocupacional, além de um impacto financeiro para as empresas [7].

A era atual é caracterizada por indústrias cada vez mais digitais, integrando pilares tecnológicos já estabelecidos pela Indústria 4.0 como, Cyber-physical System (CPS), Internet of Things (IoT) e Inteligência Artificial (IA) [8]. Esta transformação digital não só otimiza processos, como também oferece novas e poderosas ferramentas para abordar desafios persistentes, como é o caso da saúde e segurança dos trabalhadores.

Neste contexto, e olhando para o impacto significativo das LMERT, a filosofia da Indústria 5.0 aponta para um caminho claro de utilizar a tecnologia para promover uma maior e mais saudável integração entre homem e máquina.

O desafio reside em como aplicar estas tecnologias para monitorizar os fatores de risco ergonómico, como posturas inadequadas ou movimentos repetitivos, de forma precisa e contínua.

Tradicionalmente, a análise ergonómica depende de métodos de observação direta ou da utilização de sensores “vestíveis” (*wearables*) acoplados ao sistema humano [9]. Contudo, estas abordagens podem ser problemáticas, a observação direta é pontual e subjetiva, enquanto os sensores físicos podem ser invasivos, limitar a movimentação natural do trabalhador ou, paradoxalmente, alterar a própria postura que se pretende medir.

Como alternativa, os sistemas não invasivos baseados em Visão Computacional (VC) surgem como uma solução promissora. A aplicação de algoritmos de IA sobre frames de imagens de vídeo permite analisar a postura e os movimentos humanos em tempo real, sem qualquer contacto físico, alinhando-se perfeitamente com os objetivos de um sistema centrado no ser humano [10].

É neste contexto que a presente investigação se insere. A solução proposta baseia-se na aplicação de IA em conjunto com VC, permitindo o desenvolvimento de um dispositivo de baixo custo utilizando fundamentalmente dois recursos, uma câmara e um algoritmo

na linguagem de programação *Python*.

1.2 Objetivos

Esta dissertação tem como objetivo a criação de um sistema de monitorização da ergonomia do trabalhador de forma não invasiva com o auxílio de IA. Captando a imagem por uma câmara de profundidade, processando as informações em tempo real, e através de uma interface interativa, será possível apresentar os momentos de má postura e um resumo sobre a jornada de trabalho. A fim de auxiliar a verificação ergonómica no trabalho, a saída do sistema será avaliada por uma especialista.

1.3 Estrutura do Documento

Esta tese está estruturada em Introdução, Estado da Arte, Arquitetura e Implementação do Sistema, Validação e Discussão e finaliza com a Conclusão.

O Estado da Arte, **Capítulo 2**, inicia com a contextualização do problema no âmbito da Indústria 5.0, focando no bem-estar humano, na ergonomia e na análise dos distúrbios músculo-esqueléticos, revendo ainda as ferramentas de avaliação tradicionais. De seguida, a revisão foca-se na solução tecnológica, explorando a VC e os modelos de estimação de pose humana. Finalmente, é realizada uma análise de trabalhos relacionados que posicionam a presente investigação.

A Arquitetura e Implementação do Sistema, **Capítulo 3**, apresenta a estrutura técnica construída, detalhando os componentes de *hardware* e *software* selecionados. De seguida, é explicado o *pipeline* de processamento de imagem, desde a captura até ao cálculo dos ângulos posturais, e como o fluxo de dados é gerido. O capítulo conclui com a apresentação da interface gráfica final destinada ao utilizador.

Durante a Validação e Discussão, **Capítulo 4**, são apresentados algoritmos auxiliares ao sistema que surgem a medida que testes foram feitos, além de apresentar o funcionamento dos filtros de suavização e a validação do sistema junto a uma profissional na área

da ergonomia do trabalho.

Por fim, a Conclusão, **Capítulo 5**, apresenta a avaliação sobre o cumprimento dos objetivos traçados inicialmente e propõe possíveis expansões em trabalhos futuros.

Capítulo 2

Estado da Arte

Um sistema de análise ergonômica é um dispositivo computacional capaz de receber uma informação do mundo físico em forma de imagem, e ser capaz de avaliar esta imagem e os dados que ela possui, retornando os resultados obtidos sobre ela. Para tal funcionamento, faz-se necessário a utilização de tecnologias que operem em harmonia como VC e IA, complementando-se e cooperando.

Este capítulo irá introduzir os temas explorados durante a dissertação como a Indústria 5.0, ergonomia, ferramentas de avaliação ergonômica e VC.

2.1 Indústria 5.0

A indústria 5.0 é a evolução natural e complementar do paradigma da Indústria 4.0, propondo um modelo industrial que prioriza a resiliência, a sustentabilidade e, de forma central, o bem-estar humano, mantendo as características principais da antiga revolução como a automação e eficiência [1]. Este novo modelo reconhece que a automatização puramente tecnológica pode levar a sistemas de produções frágeis, suscetíveis a rupturas e negligentes ao papel fundamental do trabalhador em resolução de problemas complexos [11]. A nova fase progride, portanto, de uma visão de automação total para um modelo de cooperação entre humanos e máquinas.

O pilar humano-cêntrico (*human-centricity*) é a principal distinção da Indústria 5.0 em

comparação com a Indústria 4.0. Conforme analisado por Nahavandi [11], esta abordagem vai além da simples segurança física do trabalhador, ela visa potencializar as capacidades humanas, criando ambientes de trabalho onde a tecnologia serve como uma ferramenta para aumentar a cognição, a criatividade e o conforto do indivíduo. A meta não é substituir o humano, mas sim auxiliá-lo através de sistemas colaborativos.

A materialização deste conceito depende de um conjunto de tecnologias habilitadoras, tais como a IA, IoT, a robótica colaborativa (*cobots*) e o *Digital Twin* (DT) como as tecnologias principais para a implementação da Indústria 5.0 [12]. Estas ferramentas permitem a criação de sistemas inteligentes e adaptativos que monitorizam o ambiente de produção em tempo real, otimizam processos e, crucialmente, podem ser utilizadas para monitorizar o estado do próprio trabalhador. Dentro deste ecossistema tecnológico, a capacidade de criar representações virtuais e dinâmicas de sistemas físicos abre novas fronteiras para a análise e otimização do bem-estar humano.

Para que uma indústria seja realmente centrado no ser humano, a garantia da saúde e do conforto físico do colaborador em longo-prazo é um pré-requisito. A identificação e a gestão de fatores de risco, que possam conduzir a doenças crónicas, passam pela aplicação vital e prioritária das tecnologias da Indústria 5.0 [4], [11], [13].

2.2 Ergonomia

A ergonomia é o estudo sobre a adaptação do trabalho, das ferramentas e do ambiente às características humanas. O objetivo fundamental é garantir que as tarefas e os processos do trabalho sejam projetados para otimizar o bem-estar e a segurança das pessoas, melhorando também a eficiência [13]. No contexto da Indústria 5.0, a ergonomia torna-se um dos pilares essenciais para assegurar a saúde e a sustentabilidade da força de trabalho a longo prazo.

2.2.1 Distúrbios Músculo-Esqueléticos

Um dos principais focos da ergonomia, é a prevenção de LMERT. Estas lesões correspondem a uma série de problemas de saúde que afetam músculos, articulações, nervos e tendões, sendo frequentemente causados ou agravados pelas condições de trabalho [14].

Segundo os autores Costa e Vieira existe uma forte ligação entre o desenvolvimento de LMERT e a exposição a fatores de risco no local de trabalho [15]. Durante a jornada de trabalho, os principais riscos incluem a manutenção de posturas incorretas ou estáticas por períodos prolongados e o design inadequado das estações de trabalho [16]. O comportamento sedentário, comum neste tipo de ambiente, agrava o problema, contribuindo para dores crônicas no pescoço, ombros e costas [17].

O impacto socioeconómico que os LMERTs têm é profundo e representa um peso significativo tanto para as empresas como para os sistemas de saúde pública. Na União Europeia, os Distúrbios Músculo-Esqueléticos (DMEs) são a queixa de saúde relacionada com o trabalho mais comum, atingindo aproximadamente três em cada cinco trabalhadores [18]. A EU-OSHA reporta que estes distúrbios são também a principal causa de faltas no trabalho, conduzindo a perdas de produtividade que custam milhões de euros anualmente às economias europeias [18]. O cenário no Brasil não é diferente, as LMERTs continuam a ser as principais causas de doenças ocupacionais notificadas no país, sendo responsáveis por uma parcela significativa dos benefícios de auxílio-doença concedidos pela Previdência Social [19]. Esta realidade não só atinge financeiramente o sistema público e as empresas, mas também gera um grave impacto social, levando à diminuição da qualidade de vida e, em muitos casos, à incapacidade laboral precoce ou permanente dos trabalhadores. A alta prevalência e os custos diretos e indiretos associados a estas condições fazem da sua prevenção um objetivo prioritário.

2.2.2 Ferramentas de Avaliação Ergonómica

Para medir o risco postural e ajudar a definir ações corretivas, é necessário buscar na literatura o que configura como postura prejudicial à saúde. Segundo [20], para uma boa

postura é necessário que o corpo repouse sobre uma postura simétrica, configurando um bom equilíbrio muscular e um funcionamento harmonioso dos sistemas ósseo e articular. Distribuindo o peso corporal de maneira uniforme, minimizando o stresse sobre as articulações, ligamentos e músculos [20].

Para além da simetria, é necessário buscar quais ângulos de movimentos corporais configuram como má postura no trabalho para servir como limiar. Para isso, existem métodos de avaliação baseados na observação, entre os mais conhecidos e utilizados estão o *Rapid Upper Limb Assessment* (RULA) [21], *Rapid Entire Body Assessment* (REBA) [22] e o *Ovaco Working Posture Analysis* (OWAS) [23]. Estes métodos funcionam com um observador profissional de ergonomia, avaliando os ângulos corporais desta pessoa e, avaliando conforme cada modelo, se a postura é prejudicial ou não à saúde.

O método RULA foi criado por McAtamney e Corlett em 1993 como uma forma de avaliar rapidamente os riscos de lesões nos membros superiores, tronco e pescoço em tarefas nas quais as pessoas realizam movimentos repetitivos ou posturas forçadas [21]. O método fornece uma pontuação única de classificação de postura, calculando o risco com uma pontuação de 1 a 7 atribuída pelo observador, um *checklist* é preenchido conforme os ângulos observados e o *score* final é agrupado dentro de um dos 4 níveis de ação, como mostra a Tabela 2.1.

Tabela 2.1: Níveis de ação e pontuação do *score* RULA. Adaptado de [21].

Nível de ação	Score	Nível de risco	Ação recomendada
1	1 ou 2	Aceitável	Postura aceitável. Nenhuma ação é necessária.
2	3 ou 4	Baixo risco	Investigação adicional; mudanças podem ser necessárias.
3	5 ou 6	Médio risco	Investigação e mudanças necessárias em breve.
4	7	Risco muito elevado	Investigação e mudanças necessárias imediatamente.

De forma semelhante, o método REBA avalia o risco postural, mas com um foco no corpo inteiro, uma pontuação final é gerada a fim de entregar uma indicação de urgência e nível de risco sobre a postura e a ação que deve ser tomada [22].

O método OWAS é o método mais antigo entre eles, foi desenvolvido em 1997 e tem

como objetivo classificar de forma rápida e simples as posturas de trabalho mais comuns e prejudiciais em ambientes industriais. A classificação ocorre através de um código de quatro dígitos que representa a posição das costas, dos braços, das pernas e o nível de carga aplicada. Este código é então mapeado para uma de quatro categorias de ação, que indicam a urgência e necessidade de uma intervenção ergonómica [23].

A Tabela 2.2 traz um panorama geral das ferramentas de avaliação ergonómica citadas. Mostrando quantas posturas classificadas por grupo muscular cada uma tem, quais parâmetros são avaliados e os níveis de risco.

Tabela 2.2: Critério de classificação das ferramentas ergonómicas. Adaptado de [24].

Ferramenta	Postura*	Parâmetro	Nível de risco
OWAS	Costas (4), braços (3) e pernas (7)	Força/carga	4
RULA	Tronco (6), braço (6), antebraço (3), pernas (2), pescoço (6) e pulso (4, +2)	Força/carga, uso muscular	4
REBA	Tronco (5), braço (6), antebraço (2), pernas (4), pescoço (3) e pulso (3)	Força/carga, atividade	5

* O número entre parênteses refere-se ao número de posturas classificadas em cada ferramenta de avaliação.

É possível analisar que, para o grupo muscular tronco e pescoço, o método RULA possui mais posturas classificadas (Tronco 6 e pescoço 6) em comparação com as demais ferramentas, OWAS possui 4 classificações para as costas e REBA possui 5 para tronco e 3 para pescoço.

2.3 Visão Computacional e os Modelos de Estimação de Pose Humana

A VC é um campo científico interdisciplinar que visa ajudar os computadores a obter uma compreensão de alto nível a partir de imagens ou vídeos digitais. Para além do mero processamento de imagem, que se foca na manipulação ou melhoramento de imagens, a

VC tem como objetivo extrair, analisar e compreender o conteúdo visual para permitir a tomada de decisões [25].

Para determinados sistemas, nem sempre uma imagem bidimensional é suficiente para coletar todos os dados. Para isso, existem câmaras de profundidade capazes de coletar distâncias dentro do campo de visão do obturador, por exemplo a Microsoft Azure Kinect [26] e a Intel Realsense D415 [27], permitindo sistemas de análise mais complexos. Como a criação de um *Intelligent Personal Assistant* (IPA), capaz de auxiliar um operador na convecção de peças utilizando uma câmara de profundidade [28].

A VC é utilizada mundialmente desempenhando as mais diversas funções, como os carros autónomos [29], o reconhecimento facial em telemóveis [30], os sistema de controle de qualidade em linhas de produção [31], a análise de imagens de satélite para a agricultura de precisão [32] e de inúmeras inovações na área da saúde, particularmente na análise de imagens médicas [33]. Dentro deste vasto campo, surge também a *Human Pose Estimation* (HPE), que consiste em detetar a posição e orientação do corpo humano através da localização das suas principais articulações ou pontos-chave (*keypoints*).

Com a revolução do *Deep Learning*, diversos modelos de HPE foram desenvolvidos, cada um com diferentes arquiteturas e compromissos entre precisão e velocidade [34]. *OpenPose* é frequentemente citado como um trabalho seminal e um *benchmark* académico importante. Desenvolvido pela *Carnegie Mellon University*, a sua abordagem *bottom-up* permite-lhe detetar múltiplas pessoas numa cena com alta robustez, mas a sua complexidade de instalação e forte dependência de GPUs NVIDIA tornam-no menos acessível para aplicações de consumo em tempo real [35].

Em contraste, o *MediaPipe Pose*, desenvolvido pela Google [36], foi projetado desde o início para performance em tempo real numa vasta gama de dispositivos, incluindo CPUs e telemóveis. Baseado na arquitetura *BlazePose*, ele utiliza uma abordagem de deteção e rastreamento que o torna extremamente rápido. Além dos 33 pontos-chave 3D, ele oferece funcionalidades únicas e integradas, como a máscara de segmentação corporal, que são de grande valor para análises contextuais [37].

A família de modelos (*You Only Look Once* (YOLO)) é reconhecida pela sua velocidade

e precisão em detecção de objetos “*single-shot*”. As versões mais recentes, como o YOLOv8-Pose, estendem esta filosofia para a estimação de pose. Ao contrário do MediaPipe, que por padrão rastreia uma única pessoa, a abordagem do YOLO detecta todas as pessoas no *frame* a cada instante, oferecendo uma alternativa robusta para cenários que exigem a reavaliação contínua da cena [38], [39].

Um dos sistemas que popularizou a utilização de câmaras de profundidade em investigações de análise de movimento foi o Microsoft Kinect, introduzido originalmente em 2010 [40]. O sensor combinava uma câmara RGB, um projetor infravermelho e um detetor de profundidade baseado em luz estruturada, permitindo a aquisição de mapas 3D a baixo custo. O seu *Software Development Kit* (SDK) fornecia ainda um modelo integrado de rastreamento esquelético, capaz de identificar automaticamente as articulações de uma pessoa em tempo real. Estas características tornaram o Kinect amplamente utilizado em aplicações académicas e industriais, incluindo reabilitação, interação homem-máquina e análise de movimento, estabelecendo as bases para muitos sistemas modernos de HPE e visão computacional.

2.4 Análise Ergonómica com Visão Computacional

Para o desenvolvimento do sistema, foi realizada a busca sobre sistemas de análise ergonómica utilizando visão computacional (HPE) e o método de avaliação RULA. A Tabela 2.3 apresenta um resumo sobre as principais características que se relacionam com o presente estudo.

Tabela 2.3: Estudos sobre sistemas de análise ergonómica.

Referência	Câmara 3D	Tempo real	HPE	Ferramenta
[41]	✓	✓	Kinect SDK	RULA
[42]	✓	✓	Kinect SDK	RULA
[43]		✓	OpenPose	RULA
[44]		✓	MediaPipe	RULA

O [41] apresenta o K2RULA, um sistema semi-automático para avaliação ergonómica. Baseado na câmara de profundidade Microsoft Kinect V2, o sistema utiliza os dados do esqueleto 3D fornecido pela câmara, e aplica o método RULA para estimar o risco ergonómico. Com uma deteção em tempo real, o sistema teve sua validação ao comparar o *score* gerado pelo K2RULA com os obtidos por um sistema ótico de captura e movimento (SMART-DX 5000), com um software comercial (Jack-TAT) e um avaliador humano especialista em ergonomia.

Também com o uso do Microsoft Kinect, [42] propõe um sistema para operações de montagem industrial. O autor utiliza das coordenadas fornecidas pela câmara e aplica o método RULA. Este estudo é focado em demonstrar a facilidade e a precisão que a tecnologia disponibiliza, propondo uma estrutura modular que pode ser estendida para outros métodos, como o REBA.

O estudo realizado por Li et al. traz um método utilizando aprendizado profundo (*Deep Learning*) para estimar o risco de DME. O sistema é composto por dois modelos de rede neural profunda: um detetor de pose 2D baseado no OpenPose; um estimador de RULA que recebe as coordenadas dos pontos e calcula o nível de ação. O sistema avalia em tempo real com uma acurácia de 93% e 29 quadros por segundo, utilizando uma câmara RGB convencional e uma abordagem não invasiva [43].

Já o [44] propõe um sistema baseado em IA com o objetivo de identificar DME em plantações de arroz, onde os trabalhadores estão sujeitos a posturas repetitivas e prolongadas. A análise é realizada em tempo real utilizando o modelo de HPE *MediaPipe*. Os resultados demonstram que o sistema baseado em IA é mais rápido que métodos manuais tradicionais, oferecendo uma alternativa viável para a automação da avaliação ergonómica [44].

Estes estudos implementam uma análise ergonómica utilizando a visão computacional em diferentes cenários, com foco predominante no cálculo automático do *score* RULA completo. O Microsoft Kinect foi amplamente adotado devido ao seu SDK robusto [41], [42], porém a sua descontinuação pela Microsoft em 2023 compromete a aplicabilidade futura desta tecnologia. O *MediaPipe* é reconhecido como solução eficiente e mais rápida que

métodos tradicionais [44], contudo, observa-se uma lacuna na integração deste *framework* com sensores de profundidade modernos (RGB-D) para monitorização ergonómica em contexto industrial.

Adicionalmente, os trabalhos analisados adotam uma abordagem focada no cálculo do *score* RULA após a postura ter sido adotada, sem explorar sistemas de monitorização preventiva contínua que utilizem os ângulos de referência do método RULA como limiares para alerta em tempo real. Esta lacuna é particularmente relevante no contexto da Indústria 5.0, que prioriza o bem-estar do trabalhador através de sistemas *human-centric* capazes de fornecer *feedback* imediato e acionável, permitindo a correção postural antes do estabelecimento de padrões prejudiciais.

Desta forma, identifica-se a necessidade de desenvolver um sistema que utilize tecnologias modernas comercialmente disponíveis (RealSense D415) e *frameworks* de IA de código aberto (MediaPipe), implemente monitorização preventiva baseada em limiares angulares derivados do método RULA, ao invés de cálculo complexo de *score* completo, e forneça *feedback* em tempo real ao operador para correção postural imediata.

2.5 Resumo

Este capítulo discutiu e apresentou as tecnologias e conceitos que serão abordados ao longo do estudo. Também definiu o escopo da pesquisa, um sistema não invasivo de análise ergonómica em tempo real, que utilizará como referência os ângulos determinados no método RULA assim buscando avaliar quando o usuário está em postura prejudicial a sua saúde, evitando uma LMERT.

Capítulo 3

Arquitetura e Implementação do Sistema

Este capítulo explora a arquitetura e a implementação do sistema, detalhando cada componente e processo, desde a captação da imagem até o retorno ao usuário.

3.1 Sistema de Análise Ergonómica

O sistema de análise ergonómica é constituído por duas componentes principais: *hardware* e *software*. A componente de hardware é responsável pela captação dos dados, imagem RGB, profundidade e informação corporal extraída, bem como pela apresentação visual da interface ao utilizador. Já a componente de software processa esses dados, reconstrói a postura corporal em três dimensões e determina, em tempo real, o estado postural do utilizador.

A Figura 3.1 ilustra o funcionamento geral do sistema, destacando os módulos principais, separados por linhas pontilhadas, e o fluxo de dados que circula entre eles. Este diagrama oferece uma visão global do processo, desde a captura da informação pelo sensor até à apresentação dos resultados no ecrã.

Importante salientar que o objetivo deste sistema não é implementar o método RULA na sua totalidade, o que seria inviável devido à oclusão frequente de membros como

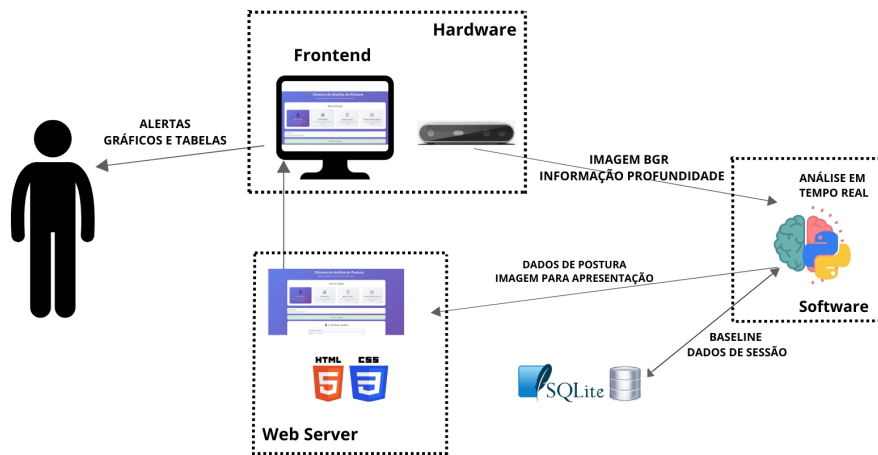


Figura 3.1: Diagrama do Sistema com Fluxo de Dados.

braços, punhos e pernas quando observados pela câmara frontal. Em vez disso, o sistema concentra-se na deteção de desvios posturais específicos do tronco e da cabeça, utilizando os limiares ergonómicos do RULA apenas como referência conceptual.

A lógica central consiste em identificar quando o utilizador ultrapassa um limite definido entre “boa postura” e “má postura”, contabilizando tanto a frequência como a duração desses eventos. Considera-se “má postura” qualquer desvio como inclinação, rotação ou assimetria que exceda uma margem pré-estabelecida em relação à postura neutra do utilizador obtida durante a calibração.

As métricas posturais monitorizadas encontram-se ilustradas na Figura 3.2. Entre elas incluem-se as inclinações frontais (pitch) da cabeça e do tronco, a inclinação lateral (roll) do tronco e a rotação (yaw) do tronco. Adicionalmente, são recolhidas duas métricas lineares não representadas na figura: a elevação média dos ombros e a assimetria entre ombro esquerdo e direito.

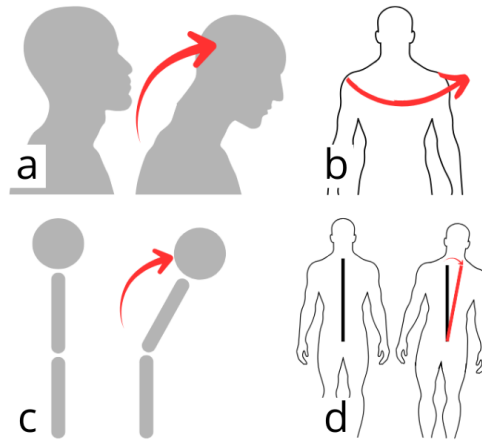


Figura 3.2: Ângulos monitorizados pelo sistema: (a) Inclinação frontal da cabeça; (b) Rotação lateral do tronco; (c) Inclinação frontal do tronco; (d) Inclinação lateral do tronco.

3.1.1 Hardware

A aquisição da imagem de profundidade é realizada pela câmara Intel RealSense D415, que opera com base no princípio da estereoscopia ativa de infravermelhos. O seu funcionamento consiste na emissão de um padrão de luz infravermelha sobre a cena, cujo reflexo é captado por um par de sensores de imagem posicionados em ângulos distintos. Um processador de visão dedicado analisa a disparidade (a diferença de posição dos pontos entre as duas imagens) e, através da triangulação, calcula a distância para cada ponto, gerando um mapa de profundidade detalhado. A interação com o hardware é facilitada pelo fabricante através de um robusto SDK (*Software Development Kit*), o *librealsense*, que disponibiliza *wrappers* para as principais linguagens de programação, incluindo a biblioteca *pyrealsense2* para Python utilizada neste trabalho, foi utilizado também bibliotecas de auxílio como OpenCV, Numpy, sqlite3 e Flask.

Para visualização da interface do usuário é necessário um ecrã conectado a um computador, e este deve ser conectado a câmara via uma porta USB 3.0. O computador é necessário para iniciar o sistema, e hospedar o *web server* localmente, sem a necessidade de conexão com a rede. Todas os componentes são conectados via cabo, sem a necessidade de conexão *wireless*.

3.1.2 Software

A parte de *software* foi projetada para ser uma arquitetura modular de três camadas, processamento, base de dados e interface.

A camada de processamento é a responsável por receber a imagem proveniente da câmara, processar os dados de postura, realizar os cálculos dos parâmetros biomecânicos, aplicar filtros de suavização, calibrar o sistema e classificar a postura. Em suma, esta camada é a camada principal, toda a lógica de análise está aqui. O algoritmo foi codificado utilizando *Python* 3.10.11.

A base de dados *SQLite* é a camada onde são armazenadas as informações a serem utilizadas durante o processo, ou para verificação posterior de uma sessão de trabalho. Esta camada armazena os usuários e seus valores, as sessões e os dados de telemetria do usuário.

E por fim, a camada de apresentação (*Web Server*) é responsável por realizar a comunicação entre o usuário e o algoritmo de processamento, fornecendo acesso em tempo real e histórico dos dados armazenados. É a parte do sistema que torna ele interativo e acessível ao usuário.

A Figura 3.3 traz um panorama geral de como o sistema funcionará. Representando graficamente todo o processo do sistema. Começando pela inicialização do sistema e a comunicação entre *Frontend* e verificação de status primordiais para o funcionamento (conexão com a câmara, conexão *WebSocket*). Depois passamos para a calibração (caso o usuário não tenha seu perfil calibrado), o usuário deve permanecer em postura neutra (ombros simétricos, coluna ereta, cabeça olhando para frente e braços ao lado do corpo), desta forma o sistema irá durante 120 *frames* calcular o valor base do usuário (*baseline*). Uma vez que tenha o usuário devidamente calibrado, o sistema está apto para analisar sua postura em tempo real, fazendo leituras angulares com uma taxa de quadros por segundo de 30 *frames* por segundo (FPS), e projetando os ângulos lidos e o tempo de sessão na *dashboard*. Assim que desejar, a análise pode ser encerrada, salvando os dados e médias de informações desta sessão na base de dados *SQLite*.

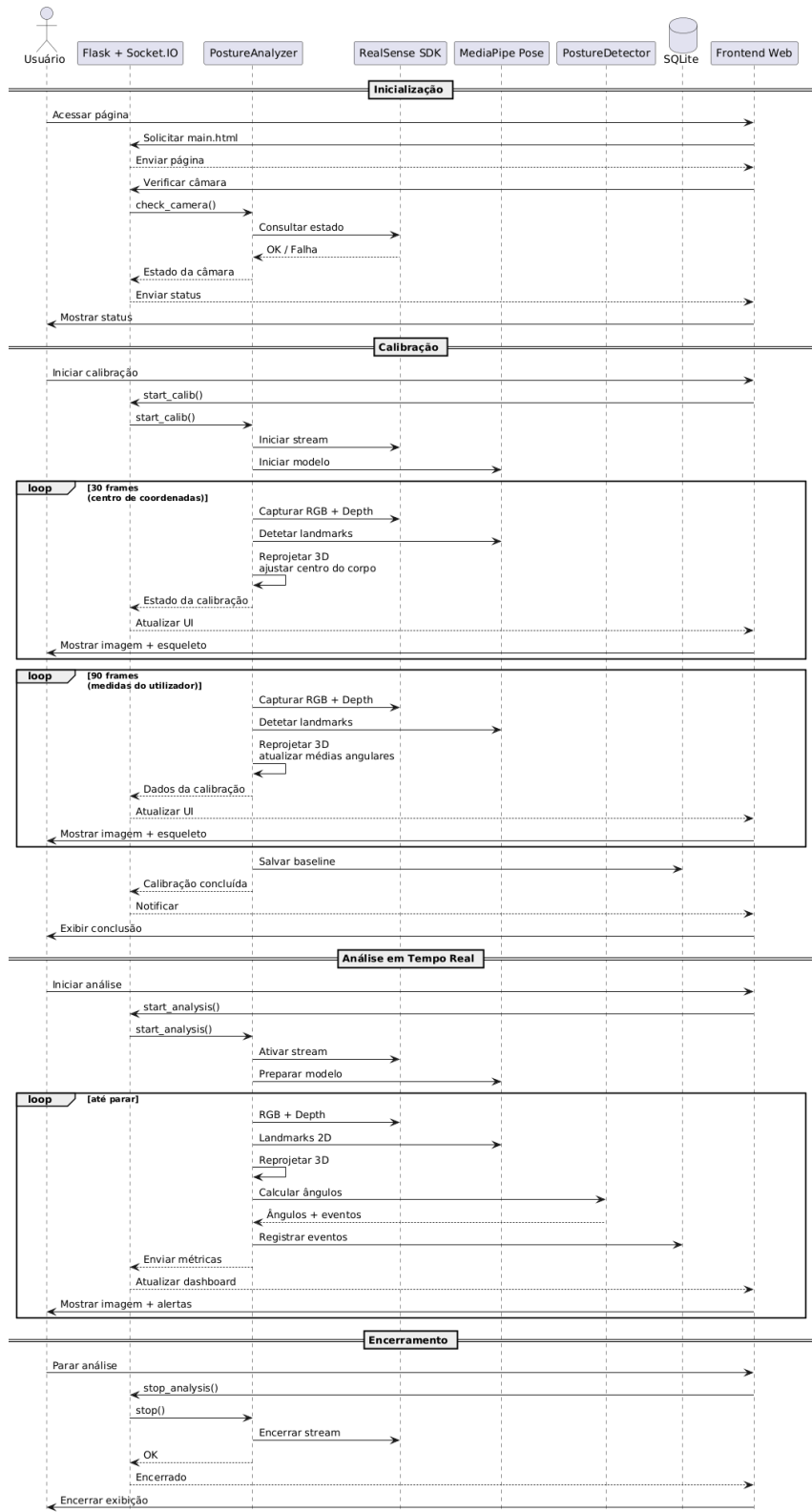


Figura 3.3: Diagrama de Sequência UML.

3.2 Detetor de Postura

O sistema foi desenvolvido utilizando a bancada de trabalho localizada no laboratório Centro de Investigação em Digitalização e Robótica Inteligente (CeDRI), dentro das dependências da Escola Superior de Tecnologia e Gestão (ESTiG) no Instituto Politécnico de Bragança (IPB). A bancada de trabalho foi desenvolvida para servir como um IPA pelo [28]. Esta bancada é equipada com um ecrã e uma câmara Intel Realsense D415, que foram aproveitados para este presente projeto.

Portanto, o sistema foi desenvolvido pensando principalmente em um usuário em pé, de frente a uma bancada e com uma câmara a sua frente e ligeiramente acima, como mostra a Figura 3.4.

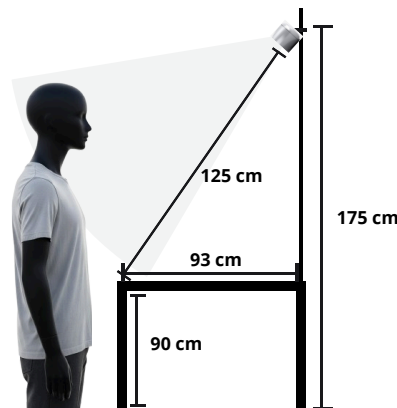


Figura 3.4: Representação Postura do Trabalhador e Captação da Imagem.

A Figura 3.5 ilustra as informações que a câmara disponibiliza, a imagem a cores (direita) e a informação sobre profundidade desta mesma imagem (esquerda), aplicado um esquema de cor do tipo Jet para representar profundidade, sendo a cor azul representando um ponto mais próximo e a vermelha um ponto mais distante.

Para a presente dissertação, o modelo de HPE escolhido foi o modelo MediaPipe. A decisão dessa escolha surge a partir da proposta de estimação de pose de cada modelo. O

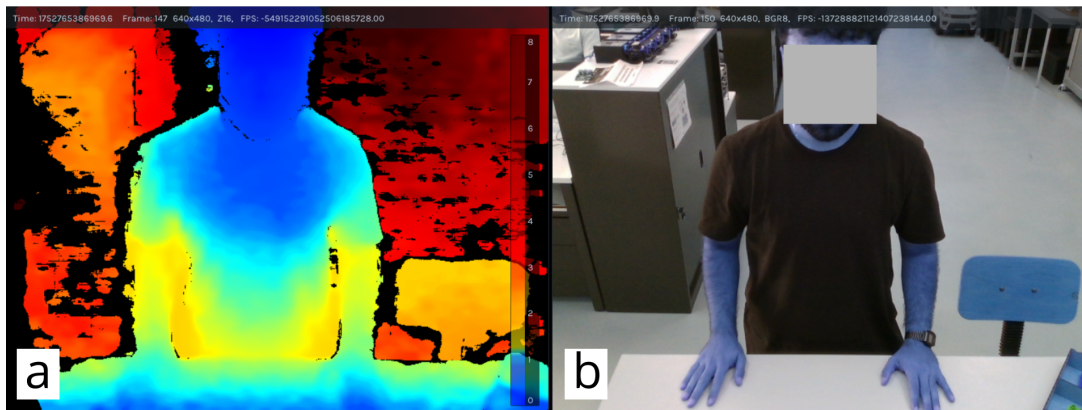


Figura 3.5: Imagem de Profundidade (a) e Imagem com Cores (b).

YOLOv8-Pose busca por varias poses em um mesmo *frame*, sendo necessário uma lógica de estimação de distância entre a detecção mais próxima do sistema para se definir o usuário, além de uma elevada necessidade de processamento para se realizar tal tarefa, tendo melhor desempenho sob a utilização de uma *Graphics Processing Unit* (GPU). Enquanto o modelo de MediaPipe deteta apenas uma pose, facilitando a lógica do sistema e diminuindo tempo de processamento, apresentando um bom desempenho somente com o uso da *Central Processing Unit* (CPU) [45].

Para transformar a informação recebida pela câmara em ângulos para serem analisados, é necessário realizar alguns processos para converter a informação de imagem para valores conhecidos. A lógica adotada segue uma linha de raciocínio convencional para análise de imagens com VC. Primeiramente fazemos o pré processamento da imagem, depois a imagem é processada e analisada e então, temos os resultados da imagem. Esta ordem de processos pode ser vista na Figura 3.6.



Figura 3.6: Representação Estrutural do Processamento de Imagem.

3.2.1 Processamento de Imagem

A câmara está configurada para transmissão de imagem em tempo real a uma resolução de 640×480 *pixels*, a uma taxa de 30 FPS. Isto significa que, a cada $\frac{1}{30}$ de segundo (≈ 33.33 ms), uma imagem contendo 307.200 *pixels* será recebida pelo algoritmo de processamento.

O processamento inicializa pela captura da imagem, com o auxílio da biblioteca *pyrealsense2*, disponibilizada pela fabricante da câmara, o *frame* de vídeo bruto recebido pela D415 é convertido de escala BGR (*Blue, Green, Red*) para a escala padrão RGB (*Red, Blue, Green*).

Com a imagem convertida para a nova escala, a informação RGB é passada para o HPE *MediaPipe Pose*. Esta tecnologia desenvolvida pela Google é capaz de receber uma imagem 2D e mapear, se estiver presente na imagem, um corpo humano com 33 marcadores corporais (*landmarks*), como mostra a Figura 3.7.

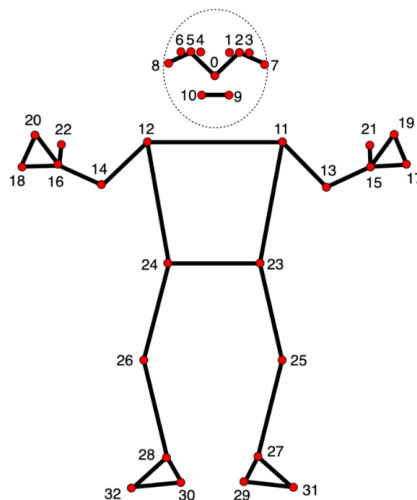


Figura 3.7: 33 Marcadores Corporais Detetados pelo *MediaPipe*. Fonte: [36]

A escolha pelo *MediaPipe Pose* foi definida após testes comparativos com outras bibliotecas de estimação de pose humana, incluindo o *YOLO-Pose*. Durante esses testes, observou-se que o *YOLO-Pose* apresentava uma flutuação ligeiramente maior na posição dos *landmarks*, mesmo em posturas estáticas, o que introduzia instabilidade indesejada para um sistema ergonômico que requer medições consistentes ao longo do tempo. Além disso, o *YOLO-Pose* demonstrou ser computacionalmente mais pesado, exigindo maior capacidade de processamento e reduzindo significativamente a taxa de fotogramas quando executado apenas em CPU. Em contraste, o *MediaPipe Pose* mostrou-se mais estável e leve, assegurando execução em tempo real sem necessidade de GPU dedicada. Para maximizar a precisão, optou-se pela utilização do modelo *Full* (*model_complexity = 1*), que apresenta melhor fidelidade na detecção dos pontos corporais e boa velocidade de processamento, e habilitou-se a segmentação permitindo a visualização da máscara corporal utilizada no processo de correção dos pontos dos quadris.

Cada ponto de referência conta com as coordenadas (x,y) normalizadas, isto é, o valor varia de 0 a 1 conforme se move pela imagem. O valor (0.0,0.0) é o canto superior esquerdo da imagem, o canto superior direito da imagem é (0.0,1.0), o inferior esquerdo é (1.0,0.0) e o inferior direito é representado por (1.0,1.0). Para encontrarmos o valor real da coordenada x,y em *pixel* é necessário multiplicar o valor normalizado pela resolução da imagem. Neste caso, chamamos a altura de 480 *pixels* como h (*height*) e a largura de 640 *pixels* w (*widht*). A Equação 3.1 apresenta o cálculo do valor real para o *pixel* x e a Equação 3.2 apresenta o cálculo para o *pixel* y.

$$x_{\text{pixel}} = x_{\text{normalizado}} \times w \quad (3.1)$$

$$y_{\text{pixel}} = y_{\text{normalizado}} \times h \quad (3.2)$$

O detetor de pose humana também retorna um valor para a coordenada z, esta é uma coordenada de profundidade relativa, também normalizada. Ele define como z=0 no centro dos quadris, valores negativos indicam que a *landmark* está mais próxima da

câmara, este valor não é uma medida em metros mas sim, um valor estimado pelo modelo. Para essa aplicação este valor é desconsiderado, pois é possível pegar a coordenada z, em metros e com uma maior precisão, pelo sensor de profundidade da câmara.

Além dessas 3 coordenadas, o modelo retorna também um valor de visibilidade (*visibility*). O valor entre 0.0 e 1.0 representa a confiança do modelo de que a *landmark* está visível e foi localizada corretamente, quanto mais próximo o valor for de 1, maior a confiança.

3.2.2 Projeção 2D para 3D

Para este projeto foi realizada a coleta dos pontos corporais apresentados na Tabela 3.1. Para garantir a qualidade dos dados, foi aplicado um filtro metodológico, aceitando-se apenas pontos com confiança (*visibility*) superior a 0.5, descartando assim estimativas de baixa precisão.

A fim de facilitar a explicação para efeito de cálculo, as coordenadas X e Y no plano do sensor 2D serão denominadas por (u,v). As coordenadas, quando passadas para o plano 3D, serão denominadas por (X,Y).

- **(X, Y, Z)**: Coordenadas tridimensionais reais do ponto no espaço.
- **(u, v)**: Coordenadas bidimensionais em pixels no plano do sensor.

Tabela 3.1: Representação dos marcadores corporais utilizados pelo *MediaPipe*.

Posição corporal	Nomenclatura 2D	Nomenclatura 3D	Número MediaPipe
Nariz	px_N	N	0
Olho esquerdo	px_LE	LE	2
Olho direito	px_RE	RE	5
Ombro esquerdo	px_LS	LS	11
Ombro direito	px_RS	RS	12
Pulso esquerdo	px_LW	LW	15
Pulso direito	px_RW	RW	16
Quadril esquerdo	px_LH	LH	23
Quadril direito	px_RH	RH	24

Após serem coletados pelo modelo, estes valores são convertidos de duas dimensões para três dimensões. Para isso, precisamos medir a profundidade em cada ponto (u,v) retornado pelo modelo de HPE.

Foi criada uma *função auxiliar* denominada “obter_ponto_3D”, esta função utiliza duas ferramentas disponibilizadas pela biblioteca *Pyrealsense2*. A primeira é a função "get_distance(u,v)", esta função irá atribuir a uma variável "d" o valor da profundidade no ponto (u,v) . Com este valor de profundidade, agora temos um ponto com as coordenadas u e v em *pixels*, e a profundidade d em metros, resultando em uma coordenada (u,v,d) .

Com esta nova informação, podemos utilizar a segunda e principal função disponibilizada pelo fabricante, a função "deproject". Esta função utiliza os parâmetros intrínsecos da câmara, distância focal e centro da lente, para calcular a posição exata daquele ponto no espaço 3D, em metros, em relação à câmara.

O calculo é baseado no modelo de câmara "pinhole". Este modelo explica por meio de semelhança de triângulos e o conceito físico câmara escura, como é calculado a distância entre um ponto no mundo real e a lente da câmara, ilustrado na Figura 3.8.

Os dados intrínsecos da câmara são únicos para cada câmara e lente e são calibrados

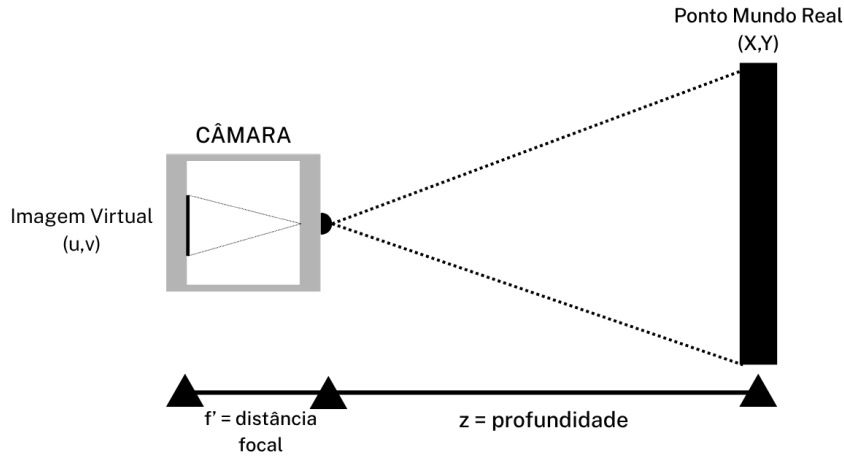


Figura 3.8: Ilustração Funcionamento Pinhole com Representação da Câmara e Objeto no Mundo Real

na fábrica. Ao obter esses dados através da biblioteca do fabricante, obtém-se as seguintes informações:

- f_x, f_y : Distâncias focais da lente, medidas em *pixels*.
- pp_x, pp_y : Coordenadas do ponto principal, ou seja, a projeção do centro óptico da lente no sensor.

Estes valores tornam o cálculo do ponto 3D possível com as seguintes equações:

$$X = d \cdot \frac{(u - pp_x)}{f_x} \quad (3.3)$$

$$Y = d \cdot \frac{(v - pp_y)}{f_y} \quad (3.4)$$

$$Z = d \quad (3.5)$$

O cálculo das coordenadas X (Equação 3.3) e Y (Equação 3.4) utiliza os parâmetros intrínsecos da câmara para projetar o *pixel* no espaço, enquanto a coordenada Z (Equação 3.5) é obtida diretamente do mapa de profundidade.

Portanto, a localização de cada ponto de interesse no plano real já é conhecida. Com estes cálculos, será obtido um ponto com coordenadas (X, Y, Z) medidas em metros.

3.2.3 Cálculo dos Ângulos Posturais

Com os pontos tridimensionais obtidos, o próximo passo consiste na criação de vetores que representam segmentos corporais e eixos de orientação. Estes vetores são a base para o cálculo dos ângulos de interesse, que descrevem a postura do usuário em relação ao espaço tridimensional.

Primeiramente, definem-se novos pontos de referência para auxiliar nos cálculos. O centro da cabeça (H) recebe a coordenada 3D do nariz. O centro dos ombros (S) é calculado pela Equação 3.6, o centro do quadril (H_{hip}) pela Equação 3.7, e a largura dos ombros (W) é determinada pela Equação 3.8:

$$\mathbf{S} = \frac{\mathbf{LS} + \mathbf{RS}}{2} \quad (3.6)$$

$$\mathbf{H}_{hip} = \frac{\mathbf{LH} + \mathbf{RH}}{2} \quad (3.7)$$

$$W = \|\mathbf{RS} - \mathbf{LS}\| \quad (3.8)$$

A partir desses pontos, constrói-se um Sistema de Coordenadas do Corpo (BCS), composto por três vetores ortogonais e unitários centrados em S .

O eixo vertical, \hat{y}_{body} , é definido pela Equação 3.9, alinhando-se com a coluna (do quadril aos ombros). O eixo horizontal, \hat{x}_{body} , é definido pela Equação 3.10, alinhando-se com a direção dos ombros. Finalmente, o eixo \hat{z}_{body} , que representa a direção frontal, é calculado pela Equação 3.11 através do produto vetorial dos anteriores, formando um sistema de coordenadas ortogonal.

$$\hat{y}_{body} = \frac{\mathbf{S} - \mathbf{H}_{hip}}{\|\mathbf{S} - \mathbf{H}_{hip}\|} \quad (3.9)$$

$$\hat{x}_{body} = \frac{\mathbf{RS} - \mathbf{LS}}{\|\mathbf{RS} - \mathbf{LS}\|} \quad (3.10)$$

$$\hat{z}_{body} = \frac{\hat{x}_{body} \times \hat{y}_{body}}{\|\hat{x}_{body} \times \hat{y}_{body}\|} \quad (3.11)$$

Esses vetores definem os eixos do tronco, dos ombros e a direção perpendicular ao plano formado por eles, conforme ilustrado na Figura 3.9. A partir deste sistema de referência local, é possível calcular todos os ângulos posturais descritos a seguir.

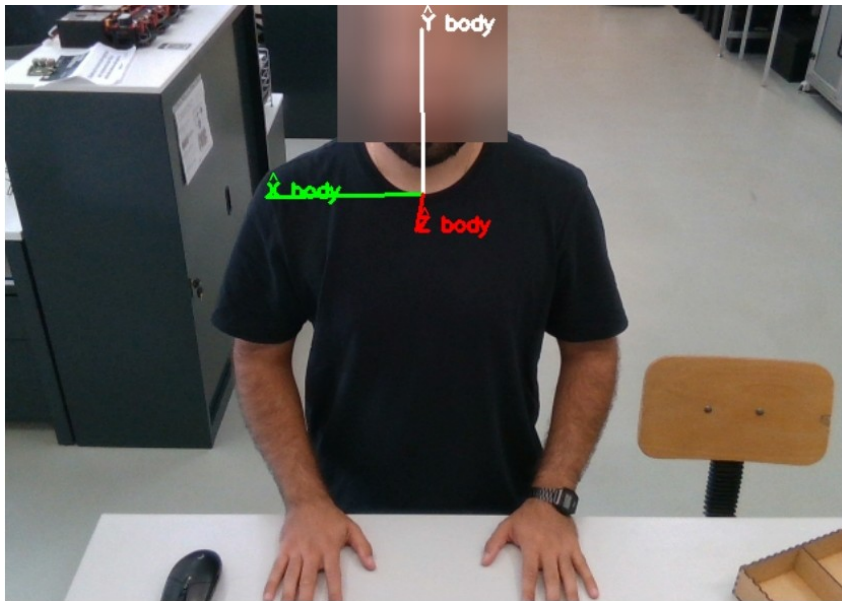


Figura 3.9: Captação dos Eixos em Relação ao Corpo do Usuário.

Planos anatômicos e eixos de referência

Com base no sistema de eixos definido, estabelecem-se os planos anatômicos:

$$\text{Plano Sagital : } (\hat{y}_{body}, \hat{z}_{body}) \quad (3.12)$$

$$\text{Plano Frontal : } (\hat{x}_{body}, \hat{y}_{body}) \quad (3.13)$$

$$\text{Plano Transversal : } (\hat{x}_{body}, \hat{z}_{body}) \quad (3.14)$$

Os ângulos posturais são obtidos por operações vetoriais, utilizando produtos escalar e vetorial, de forma a quantificar a orientação de cada segmento corporal em relação a esses planos.

Pitch da Cabeça

O *pitch* da cabeça representa a inclinação da cabeça em relação ao eixo vertical do tronco, dentro do plano sagital. Sendo \hat{v}_{head} o vetor médio entre as orelhas (direita e esquerda) e \hat{y}_{body} o eixo vertical do tronco, o ângulo é calculado conforme a Equação 3.15:

$$\theta_{pitch}^{head} = \arccos \left(\frac{\hat{v}_{head} \cdot \hat{y}_{body}}{\|\hat{v}_{head}\| \|\hat{y}_{body}\|} \right) \quad (3.15)$$

Um valor positivo de θ_{pitch}^{head} indica flexão anterior da cabeça (inclinação para frente).

Pitch do Tronco

Para cálculo da inclinação frontal do Tronco, foi criado um eixo auxiliar chamado (\hat{y}_{world}) que recebe o valor de (\hat{y}_{body}), como mostra a Equação 3.16, o (\hat{y}_{world}) é o eixo de referência calculado durante a calibração, este eixo é fixo e serve como base para cálculo dos valores instantâneos.

$$\hat{y}_{world} = \hat{y}_{body} \quad (3.16)$$

O *pitch* do tronco ou, inclinação frontal, é calculado como o ângulo entre o eixo vertical

do corpo (\hat{y}_{body}) e o eixo vertical global (\hat{y}_{world}), conforme a Equação 3.17:

$$\theta_{pitch}^{trunk} = \arccos \left(\frac{\hat{y}_{body} \cdot \hat{y}_{world}}{\|\hat{y}_{body}\| \|\hat{y}_{world}\|} \right) \quad (3.17)$$

Valores positivos representam flexão anterior do tronco.

Yaw dos Ombros

O *yaw* dos ombros corresponde à rotação do eixo dos ombros no plano transversal, refletindo a torção axial do tronco. Sendo $\hat{v}_{shoulders}$ o vetor entre os ombros direito e esquerdo, projetado no plano ($\hat{x}_{body}, \hat{z}_{body}$), o ângulo é dado pela Equação 3.18:

$$\theta_{yaw}^{shoulders} = \arctan 2 (\hat{v}_{shoulders} \cdot \hat{z}_{body}, \hat{v}_{shoulders} \cdot \hat{x}_{body}) \quad (3.18)$$

Valores positivos indicam rotação do tronco para a esquerda (vista superior).

Roll do Tronco

O *roll* do tronco indica a inclinação lateral do corpo (plano frontal). É obtido pelo ângulo entre o eixo vertical do tronco e o eixo vertical global, como descrito na Equação 3.19:

$$\theta_{roll}^{trunk} = \arctan 2 (\hat{y}_{body} \cdot \hat{x}_{world}, \hat{y}_{body} \cdot \hat{y}_{world}) \quad (3.19)$$

Valores positivos indicam inclinação para o lado direito.

Roll dos Ombros

O *roll* dos ombros é o ângulo de inclinação da linha entre os ombros em relação à horizontal, calculado pela Equação 3.20:

$$\theta_{roll}^{shoulders} = \arctan 2(y_R - y_L, x_R - x_L) \quad (3.20)$$

Este ângulo é útil para detetar desalinhamentos ou inclinações laterais.

Assimetria e Elevação dos Ombros.

Diferentemente das grandezas anteriores, as métricas de *elevação média* (EM) e *assimetria* (ED) não são ângulos, mas sim medidas lineares **adimensionais**, normalizadas pela largura efetiva dos ombros W_{eff} . Ambas são obtidas diretamente das coordenadas tridimensionais (em metros) dos ombros esquerdo e direito.

Para isso, calcula-se primeiro a altura média dos ombros (\bar{y}) no eixo vertical da câmara, conforme a Equação 3.21. A Elevação Média (EM), definida na Equação 3.22, compara a altura média atual (\bar{y}) com a altura de repouso (\bar{y}_0). Por sua vez, a Assimetria (ED), detalhada na Equação 3.23, mede a diferença de altura entre os ombros esquerdo e direito.

$$\bar{y} = \frac{1}{2}(y_{LS} + y_{RS}) \quad (3.21)$$

$$\text{EM} = \frac{\bar{y}_0 - \bar{y}}{W_{\text{eff}}} \quad (3.22)$$

$$\text{ED} = \frac{y_{LS} - y_{RS}}{W_{\text{eff}}} \quad (3.23)$$

Como o eixo Y da câmara RealSense cresce para baixo, $\text{EM} > 0$ (Equação 3.22) indica elevação dos ombros (menor valor de y), enquanto o sinal de ED (Equação 3.23) identifica o lado mais alto: $\text{ED} > 0$ significa ombro esquerdo mais baixo e $\text{ED} < 0$ ombro esquerdo mais alto. Na interface gráfica, ambas são exibidas em forma percentual, $100 \times \text{EM} \%$ e $|\text{ED}| \times 100 \%$. Enquanto os parâmetros de *pitch*, *roll* e *yaw* são baseados em ângulos no espaço tridimensional, as métricas EM e ED representam variações relativas de posição essenciais para detetar elevação de ombros e desalinhamentos laterais mesmo quando a postura global do tronco permanece estável.

Todos os ângulos calculados são suavizados ao longo do tempo, utilizando filtros de média móvel para reduzir ruídos de detecção. Os resultados são registados em banco de dados (SQLite) e comparados com a postura de referência calibrada, permitindo a análise temporal e a geração de métricas ergonómicas consistentes.

3.2.4 Filtros de Suavização

Durante a etapa de cálculo dos ângulos posturais, observou-se a presença de ruídos característicos de sistemas de visão computacional, especialmente em momentos de oclusão parcial ou pequenas instabilidades no rastreamento dos pontos corporais. Para mitigar esses efeitos e garantir maior estabilidade temporal, foi implementado um módulo de filtragem composto por duas abordagens complementares: um filtro de Média Móvel Exponencial (EMA) e um filtro de Mediana. Além disso, um mecanismo de persistência temporal, denominado *Latch*, foi incorporado para evitar ativações falsas devido a flutuações momentâneas.

Filtro de Média Móvel Exponencial (EMA)

O primeiro filtro desenvolvido foi o de Média Móvel Exponencial (*Exponential Moving Average* (EMA)), aplicado principalmente aos ângulos de inclinação, como o *pitch* do tronco. Esse tipo de filtro é amplamente utilizado para suavizar séries temporais, reduzindo oscilações rápidas e enfatizando a tendência principal do movimento. A implementação ocorre de forma recursiva, utilizando apenas a medição mais recente e o valor filtrado anterior, conforme descrito na Equação 3.24.

$$S_t = \alpha \cdot Y_t + (1 - \alpha) \cdot S_{t-1} \quad (3.24)$$

Onde:

- S_t : valor filtrado no instante de tempo atual (t);
- Y_t : valor observado (ângulo bruto) no instante t ;
- S_{t-1} : valor filtrado no instante anterior;
- α : fator de suavização ($0 < \alpha \leq 1$), que define o peso dado às novas observações.

A principal vantagem do EMA é seu caráter adaptativo e de baixa complexidade

computacional. Ele mantém apenas uma “memória condensada” do histórico, ajustando-se gradualmente a novas variações. Assim, contribui para a remoção de ruídos de alta frequência sem introduzir atrasos perceptíveis ou distorcer o comportamento geral do movimento.

Na prática, a aplicação do EMA resultou em curvas mais estáveis e coerentes com a dinâmica esperada do operador. As derivadas temporais ($dPitch/dt$) evidenciaram uma redução aproximada de 67% na variabilidade angular, comprovando que o filtro atenua oscilações espúrias sem comprometer a resposta a movimentos reais. A Figura 3.10 ilustra o efeito visual obtido.

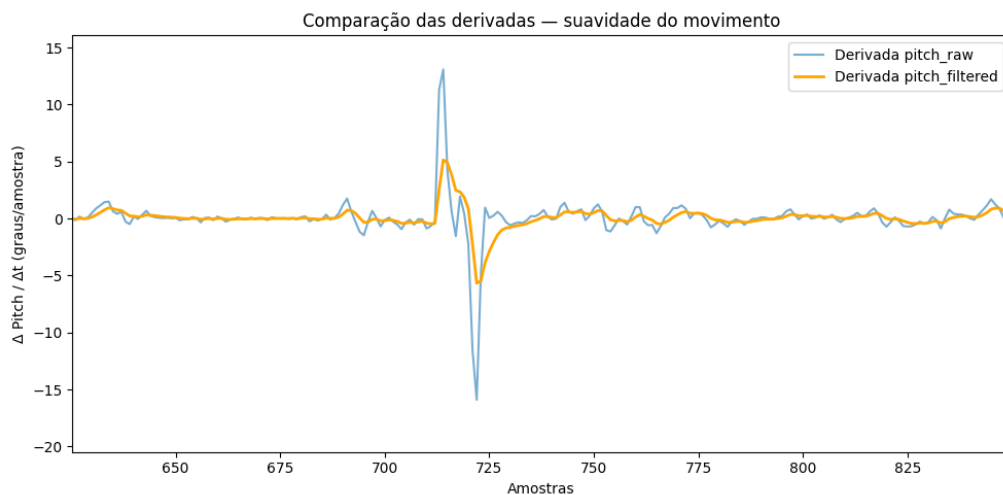


Figura 3.10: Comparação entre as Derivadas Temporais do Ângulo de *Pitch* Antes e Após a Filtragem.

Filtro de Mediana

O segundo filtro implementado foi o de Mediana, utilizado principalmente para ângulos de rotação, como o *yaw*. Diferente da média móvel, a mediana não realiza uma suavização contínua, mas é altamente eficaz na eliminação de ruídos do tipo *spike*, picos isolados que surgem devido a erros momentâneos de rastreamento ou à reinicialização dos valores angulares.

O funcionamento baseia-se em uma janela deslizante de 5 valores que armazena as últimas medições do sinal. A cada nova leitura, os valores dentro dessa janela são ordenados, e o elemento central (mediana) é escolhido como saída filtrada, descartando valores extremos. Essa abordagem confere robustez contra anomalias pontuais, preservando, contudo, transições reais de movimento.

Nos experimentos, o filtro de mediana apresentou desempenho consistente na redução de variações abruptas e ruídos pontuais observados no sinal bruto de *yaw*. A aplicação resultou em uma atenuação aproximada de 20% no desvio-padrão das derivadas (de 49,94 para 39,61), indicando maior estabilidade temporal sem comprometer a resposta a mudanças reais de orientação. A diferença média entre o sinal filtrado e o original foi de 23,09°, valor condizente com o comportamento esperado de um filtro que elimina picos isolados, mas preserva as transições naturais do movimento. Assim, a filtragem por mediana contribuiu para a continuidade e coerência dos movimentos rotacionais, garantindo robustez frente a ruídos momentâneos de rastreamento. A Figura 3.11 exemplifica esse comportamento.

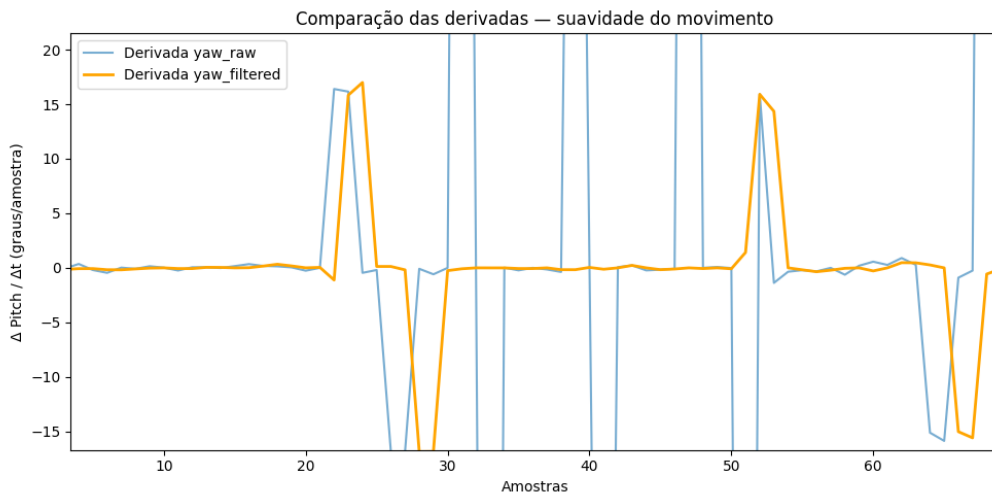


Figura 3.11: Comparação entre as Derivadas Temporais do Ângulo de *Yaw* Antes e Após a Filtragem.

Critério de Persistência Temporal (*Latch*)

Além dos filtros, foi implementado um critério de persistência temporal denominado *Latch*, responsável por garantir decisões mais confiáveis a partir de sinais potencialmente instáveis. O mecanismo impede ativações falsas provocadas por oscilações rápidas próximas aos limiares de decisão.

O funcionamento consiste em dois estágios:

- O evento de alerta é ativado somente se o sinal permanecer acima do limiar definido por um período mínimo (em número de *frames*), evitando ativações instantâneas;
- A desativação ocorre apenas quando o valor filtrado cai abaixo de um segundo limiar, entre 75% e 85% do valor original, prevenindo o “pisca” do evento caso o sinal oscile em torno do ponto de ativação.

Dessa forma, a combinação entre filtragem e persistência temporal aumenta substancialmente a confiabilidade do sistema, assegurando que apenas padrões consistentes e duradouros de movimento sejam interpretados como desvios posturais significativos.

3.2.5 Detecções Auxiliares

Durante o desenvolvimento do sistema, notou-se a necessidade de desenvolver algoritmos para auxiliar no funcionamento. Esses algoritmos servem para auxiliar o modelo de HPE a detetar melhor os pontos de interesse, e também entregar mais informações quanto a postura em pé ou sentada do usuário.

Máscara da Mesa

Durante os testes iniciais, observou-se que o modelo de estimativa de pose humana (*MediaPipe Pose*) apresentava erros recorrentes na detecção dos quadris (*hips*) sempre que o usuário estava posicionado atrás de uma mesa. Nessas situações, o modelo tendia a projetar os pontos dos quadris sobre o tampo da mesa, uma vez que parte do corpo permanecia

parcialmente ocluída na imagem de cor. Esse erro gerava deslocamentos artificiais no cálculo dos vetores corporais, distorcendo a estimativa de ângulos do tronco e da inclinação geral do corpo.

Para corrigir esse problema, foi desenvolvido um módulo de detecção automática da mesa, capaz de identificar o plano correspondente à superfície de trabalho e gerar uma máscara binária de exclusão. Essa máscara é usada pelo sistema para corrigir, em tempo real, os pontos dos quadris detetados pelo *MediaPipe*, garantindo que sejam reposicionados acima da mesa sempre que o modelo de pose os confunde com a superfície.

O processo de detecção utiliza o mapa de profundidade fornecido pela câmara Intel RealSense. O algoritmo seleciona a faixa inferior da imagem, onde normalmente a mesa está localizada, e a partir da distribuição dos valores de profundidade, ajusta o plano dominante por meio do método RANSAC (*Random Sample Consensus*). O plano encontrado é convertido em uma máscara binária, na qual os pixels pertencentes à mesa são marcados em branco. Em seguida, são aplicadas operações morfológicas de fechamento e suavização para eliminar ruídos e preencher pequenas falhas, além de uma expansão vertical de cerca de 3% da imagem, garantindo uma margem de segurança contra pequenas variações de profundidade.

Durante a análise postural, essa máscara é usada para validar a posição dos quadris esquerdo e direito. Caso algum desses pontos seja identificado dentro da área da mesa, o sistema desloca-o alguns *pixels* acima, de modo a simular sua posição real no corpo. Esse procedimento corrige as falhas do modelo de pose e assegura que os vetores corporais sejam calculados de forma coerente com a anatomia do usuário.

Em síntese, o módulo de máscara da mesa atua como um filtro espacial que complementa o algoritmo de estimativa de pose, evitando que o modelo interprete superfícies do ambiente como partes do corpo humano. Essa abordagem melhora significativamente a estabilidade e a precisão das medições posturais em contextos reais de uso, como escritórios e bancadas de trabalho.

A Figura 3.12 mostra as detecções dos quadris (*hips*) antes da aplicação do filtro (a) e depois da aplicação do filtro (b).

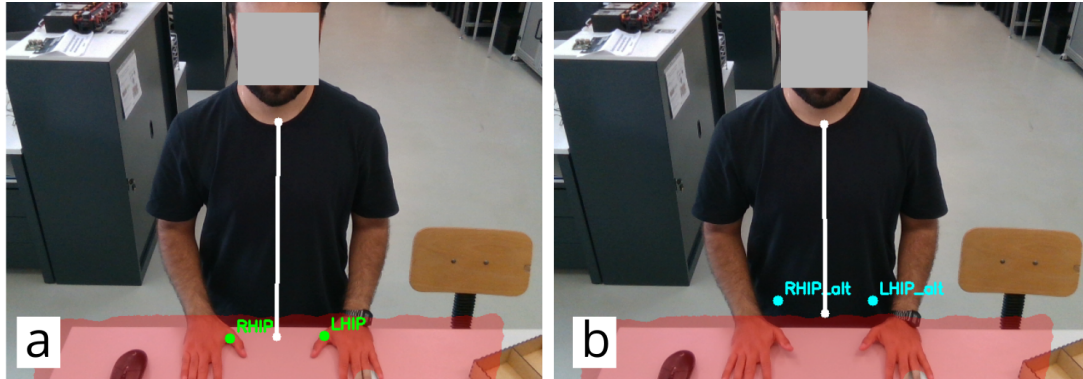


Figura 3.12: (a) Quadris Detetados Sobre a Mesa. (b) Quadris Corrigidos Para Fora da Mesa.

Detecção Sentado ou Em Pé

A análise postural é realizada apenas quando o sistema reconhece que o usuário está na posição em pé. Para isso, implementou-se um mecanismo automático que distingue entre as posturas sentado e em pé, garantindo que os cálculos de ângulos e métricas sejam realizados apenas em condições adequadas de observação.

A decisão é tomada com base em duas pistas geométricas derivadas das coordenadas tridimensionais captadas pela câmara RealSense: a variação vertical dos ombros e o avanço do tronco em relação à calibração. A postura é classificada como *sentado* apenas quando ambas as pistas excedem seus limiares de decisão, como detalhado a seguir.

1. **Queda dos ombros ($dropW$).** Mede o quanto os ombros se deslocaram verticalmente em relação à posição de referência obtida na calibração (em pé). Essa diferença é normalizada pela largura efetiva dos ombros W_{eff} , conforme a Equação 3.25, que utiliza a altura média \bar{y} calculada na Equação 3.26:

$$dropW = \frac{\bar{y} - \bar{y}_0}{W_{\text{eff}}} \quad (3.25)$$

$$\bar{y} = \frac{1}{2}(y_{LS} + y_{RS}). \quad (3.26)$$

Quando uma pessoa se senta, os ombros naturalmente ficam mais baixos em relação

à câmara, tornando $dropW$ positivo.

2. **Avanço do tronco (dz).** Mede a variação de profundidade do ponto torácico central (z_{chest}) em relação à posição de calibração, conforme a Equação 3.27:

$$dz = z_{\text{chest}} - z_{\text{chest}0}. \quad (3.27)$$

Ao sentar-se, o tronco tende a deslocar-se ligeiramente para a trás, aumentando a distância até a câmara.

A decisão final é positiva para a condição “sentado” somente se as duas pistas forem verdadeiras ao mesmo tempo, como mostra a Equação 3.28:

$$\text{sitting} = (\text{drop}W > \text{limiar}W) \wedge (dz > \text{limiar}Z). \quad (3.28)$$

O resultado dessa deteção controla o estado operacional do sistema. Quando o usuário é classificado como em pé, a análise ergonómica completa é ativada: o sistema calcula continuamente os ângulos posturais e as métricas lineares, armazena os resultados e atualiza os contadores de tempo de análise.

Por outro lado, quando o usuário é classificado como sentado, o sistema entra em modo de pausa. O cálculo de ângulos é temporariamente suspenso, uma medida essencial para evitar leituras incorretas e ruídos que poderiam comprometer a avaliação ergonómica, sendo contabilizado apenas o tempo de permanência sentado.

Além desta lógica de ativação, o sistema inclui uma funcionalidade de alerta para o ortostatismo prolongado. Se for detetada atividade contínua na postura em pé por um período superior a 30 minutos, o sistema emite uma recomendação de pausa. Esta intervenção visa mitigar os riscos ergonómicos associados a permanecer em pé de forma estática, como a fadiga muscular, dores lombares e problemas circulatórios nos membros inferiores [46], [47].

Dessa forma, o sistema adapta automaticamente o processamento conforme a postura global do usuário. Essa estratégia assegura que apenas os períodos efetivos de trabalho

em pé sejam analisados, enquanto as pausas ou mudanças de posição são corretamente registradas e descontadas da análise postural.

3.2.6 Definição de Limiar e Eventos

A definição dos limiares de ativação (o ponto de "má postura") baseia-se numa abordagem híbrida, que considera tanto os padrões ergonômicos (RULA) como as limitações práticas do sensor.

Primeiro, foi necessário estabelecer uma margem de tolerância mínima, associada à imprecisão inerente à estimativa de pose (HPE). Em repouso, observou-se uma flutuação natural do sistema inferior a 3°. Este valor foi adotado como o limiar mínimo de detecção, qualquer movimento abaixo disto é considerado ruído do sensor e deve ser ignorado.

Segundo, analisaram-se os limiares do RULA. O método classifica a inclinação da cabeça de 0-10° como nível 1 (aceitável) e 10-20° como nível 2 (primeiro sinal de risco). Para o tronco, o nível 2 começa apenas aos 20° de inclinação.

Ao combinar estas duas fontes, definiu-se uma estratégia de limiares diferenciada. Para a inclinação da cabeça, o limiar de 10° do RULA, que marca a transição para o nível 2, foi adotado diretamente por ser um limite claro e sensível. Já para a métrica do tronco, o limiar de 20° do RULA foi considerado demasiado permissivo para um sistema de monitorização em tempo real, uma vez que muitos desvios menores não seriam detetados. Por essa razão, adotou-se para o tronco o limiar de 3°, correspondente à margem de ruído do sensor, o que na prática significa que qualquer desvio significativo e não-trivial para fora da postura neutra é detetado. O mesmo foi adotado para as demais métricas como, rotação lateral do tronco, inclinação lateral do tronco, assimetria dos ombros e elevação dos ombros não possuem valor estimado no modelo RULA.

Para evitar que o sistema comute repetidamente entre “bom” e “mau” quando o utilizador oscila em torno do limiar, foi implementado o mecanismo de persistência *Latch*.

Isto é feito através de dois limiares distintos: um Limiar *On* (para ativar o evento de má postura) e um Limiar *Off* (para desativá-lo). O Limiar *Off* foi definido como 90%

do valor do Limiar *On*. Os valores finais para todas as métricas são apresentados na Tabela 3.2.

Tabela 3.2: Limiar de ativação (*On*) e desativação (*Off*) utilizados para cada medida.

Medida	Limiar <i>On</i>	Limiar <i>Off</i>
Inclinação frontal da cabeça (<i>pitch</i>)	10°	9°
Inclinação frontal do tronco (<i>pitch</i>)	3°	2.7°
Rotação lateral do tronco (<i>yaw</i>)	3°	2.7°
Inclinação lateral do tronco (<i>roll</i>)	3°	2.7°
Assimetria dos ombros	5 %	4.5 %
Elevação dos ombros	3 %	2.7 %

Finalmente, os limiares por si só não são suficientes para lidar com movimentos muito rápidos ou “pico” momentâneo no sinal. Para garantir que um evento só é contabilizado se a má postura for mantida, foi introduzido um mecanismo de persistência temporal.

Um evento de má postura só é validado (ativando o *Latch*) se a medição permanecer acima do Limiar *On* por um número mínimo de *frames* consecutivos. Para suavizar o sinal de entrada antes de o comparar com o limiar, foram também aplicados filtros de Média Móvel Exponencial (EMA) ou Mediana.

Os parâmetros finais de filtro e persistência, obtidos através de testes empíricos com vídeos de referência em diferentes condições, são detalhados na Tabela 3.3.

3.2.7 Estrutura Geral de Dados do Sistema

Durante a etapa de calibração, o sistema calcula e armazena os parâmetros de referência biomecânicos individuais de cada usuário. Esses dados representam a postura neutra (posição ergonômica correta) e servem como base para todos os cálculos diferenciais de ângulos e métricas posteriores. O arquivo gerado é armazenado localmente na base de dados `SQLite`, e o seu conteúdo também pode ser exportado em formato `JSON` para uso externo ou reimportação. A Tabela 3.4 apresenta os principais campos e seus significados no registo de calibração.

Tabela 3.3: Filtros aplicados e tempo de persistência adotado para cada métrica do sistema.

Medida	Filtro	Persistência [<i>frames</i>]
Inclinação frontal da cabeça (<i>pitch</i>)	EMA	10
Inclinação frontal do tronco (<i>pitch</i>)	Mediana	10
Rotação lateral do tronco (<i>yaw</i>)	Mediana	10
Inclinação lateral do tronco (<i>roll</i>)	EMA	10
Assimetria dos ombros	EMA	10
Elevação dos ombros	EMA	10

Tabela 3.4: Campos armazenados na *baseline* do usuário durante o processo de calibração.

Campo	Descrição	Tipo / Unidade
<code>timestamp</code>	Momento da calibração	Inteiro
<code>up</code>	Vetor unitário de orientação vertical do tronco	Vetor 3D (X, Y, Z)
<code>mu_pitch</code>	Ângulo médio de (<i>pitch</i>)	Float (°)
<code>mu_yaw</code>	Ângulo médio de (<i>yaw</i>)	Float (°)
<code>mu_roll</code>	Ângulo médio de (<i>roll</i>)	Float (°)
<code>trunk_pitch</code>	Inclinação média frontal do tronco	Float (°)
<code>trunk_roll</code>	Inclinação lateral média do tronco	Float (°)
<code>ybar0</code>	Posição vertical média do ponto torácico central	Float (m)
<code>w0</code>	Largura média dos ombros normalizada	Float (m)
<code>z_chest0</code>	Profundidade média do ponto torácico	Float (m)
<code>L_tronco0</code>	Comprimento relativo do tronco	Float (m)
<code>fps</code>	Taxa de quadros utilizada durante a calibração	Inteiro (FPS)

O formato serializado em JSON dos dados da *baseline* é ilustrado no Bloco de Código 3.1. Esse registo é gerado automaticamente após a estabilização dos vetores corporais e representa a postura de referência individual de cada usuário.

Bloco de Código 3.1: Exemplo real do arquivo de baseline armazenado no banco de dados.

```
{
  "timestamp": 1761323215,
  "up": [0.0011, -0.9381, -0.3464],
  "baseline": {
    "mu_pitch": 162.4891,
    "mu_yaw": -174.9370,
    "mu_roll": -0.4757,
    "ybar0": -0.1356,
    "W0": 0.3347,
    "z_chest0": 1.1378,
    "L_tronco0": 0.0,
    "trunk_pitch": -20.3281,
    "trunk_roll": -1.6051
  }, "fps": 30
}
```

A variável `up` representa o vetor unitário de orientação vertical estimado a partir do tronco, utilizado para projetar os vetores corporais no sistema de coordenadas do corpo. O dicionário `baseline` contém os parâmetros médios calculados a partir de múltiplos quadros estáveis durante a calibração (por padrão, $N = 90$ quadros). Esses valores formam o conjunto $\{\mu_i\}$ que serve de referência para o cálculo dos desvios angulares durante a análise, segundo a Equação 3.29:

$$\Delta\theta_i = \theta_i - \mu_i \quad (3.29)$$

onde θ_i representa o valor instantâneo medido em tempo real e μ_i o valor médio da baseline. Dessa forma, cada sessão de análise é automaticamente ajustada às características anatômicas e posturais individuais do usuário.

Além do registo individual de calibração (*baseline*), o sistema de ergonomia mantém uma estrutura de dados relacional que organiza todas as informações provenientes das análises, sessões e relatórios de cada utilizador. Essa estrutura é implementada em um banco de dados `SQLite`, garantindo portabilidade, baixo custo computacional e consistência entre os módulos de calibração, captura e avaliação de postura.

A Figura 3.13 apresenta o Diagrama Entidade–Relacionamento (DER) do sistema, representando as principais entidades e suas relações. Cada utilizador (`USERS`) possui um conjunto de sessões de análise (`SESSIONS`), nas quais são registados parâmetros temporais e métricas

ergonômicas. A partir de cada sessão, o sistema gera automaticamente um relatório sintético (REPORTS), que consolida os resultados e pode ser exportado em formato JSON para posterior análise estatística.

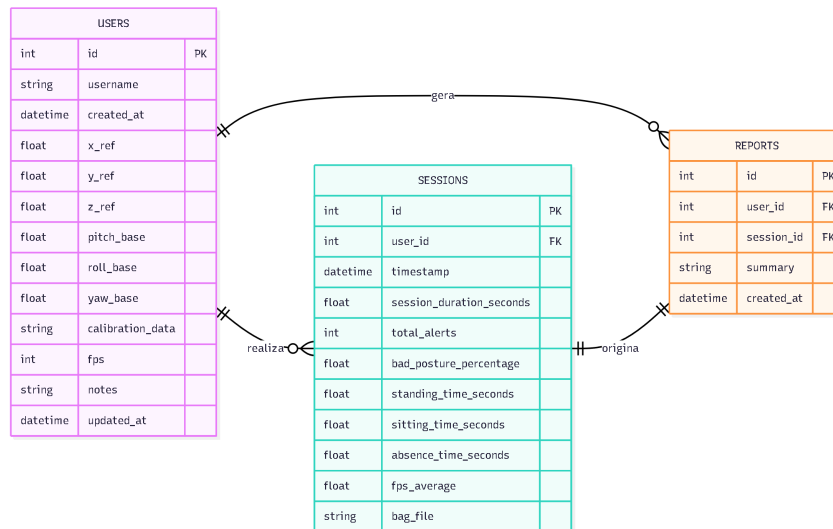


Figura 3.13: Diagrama Entidade-Relacionamento (DER) da Estrutura Geral do Sistema de Ergonomia.

Esse modelo de dados reflete a hierarquia operacional do sistema cada utilizador passa por uma etapa inicial de calibração, seguida de múltiplas sessões de monitoramento. Os dados de cada sessão são agregados em relatórios individuais, que podem posteriormente ser utilizados para análises comparativas, estatísticas ou de evolução postural ao longo do tempo.

3.3 Interface Gráfica

A interface gráfica do sistema foi desenvolvida com o objetivo de facilitar a interação entre o usuário e o módulo de processamento, permitindo o controlo completo das etapas de calibração, análise e gestão de usuários. Construída em ambiente *web* e hospedada localmente através do *framework* Flask, a interface foi implementada em *HTML*, *CSS* e *JavaScript*, garantindo compatibilidade com navegadores modernos e execução independente de conexão à internet. A comunicação em tempo real entre o servidor e o navegador é realizada via protocolo *WebSocket*, utilizando a biblioteca *Flask-SocketIO*.

A Figura 3.14 ilustra a página inicial do sistema. Nela, o Menu Principal é apresentado em



Figura 3.14: Página Inicial da Interface Gráfica do Sistema.

quatro blocos funcionais: *Criar Usuário*, *Iniciar Análise*, *Apagar Usuário* e *Verificar Últimas Leituras*. Cada bloco contém um ícone e uma breve descrição da sua função, permitindo uma navegação intuitiva e organizada. Logo abaixo, um painel de status de conexão indica o estado do sistema, com um marcador verde para o servidor ativo e outro marcador vermelho caso a câmara Intel RealSense não seja detetada.

A partir do menu principal, o usuário pode aceder à área de Criação e Calibração de Usuário, apresentada na Figura 3.15. Nessa secção, é possível registar um novo utilizador, introduzindo o seu nome e iniciando o processo de calibração. Três botões principais controlam o fluxo dessa etapa: *Iniciar Calibração*, *Parar Calibração* e *Visualizar Câmara*. Abaixo dos controlos, um bloco de instruções passo a passo orienta o operador quanto à posição correta, sequência de ações e tempo de imobilidade necessário para o registo dos dados, reduzindo erros durante a calibração.

Após o processo de calibração, o sistema permite iniciar a Análise em Tempo Real, na qual as medições são apresentadas em blocos informativos atualizados continuamente. Essa secção foi dividida em duas visualizações distintas: *Visão Operador* e *Visão Gestor*, alternáveis por um botão na parte superior da interface.

Na Visão Operador, mostrada na Figura 3.16, a interface apresenta um painel simplificado com informações essenciais para o acompanhamento da sessão, como o tempo total de análise, o tempo em descanso e a percentagem do tempo que o usuário teve má postura. Blocos coloridos facilitam a leitura rápida do estado postural e do tempo decorrido, permitindo uma

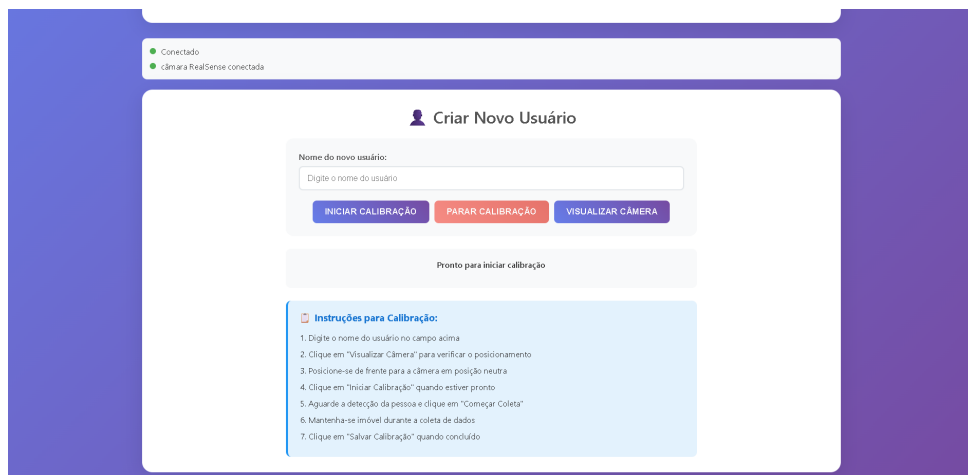


Figura 3.15: Tela de Criação e Calibração de Novo Usuário.

monitorização contínua durante o trabalho.

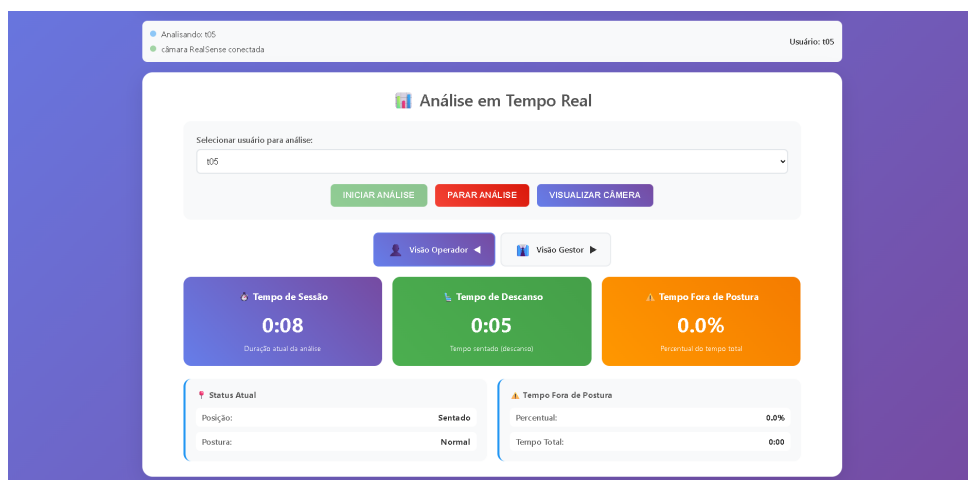


Figura 3.16: Modo de Análise em Tempo Real — Visão Operador.

Já a Visão Gestor, representada na Figura 3.17, oferece uma leitura mais detalhada dos parâmetros de postura e telemetria. Os ângulos de *pitch* da cabeça, Rotação dos ombros e inclinação do tronco são exibidos em blocos dedicados, juntamente com as métricas dos ombros (elevação, assimetria e largura). Outro conjunto de blocos apresenta dados de telemetria avançada, como duração da sessão, tempo em pé e tempo em má postura. Um painel de alertas detalhados lista as possíveis situações de risco ergonômico, como inclinação excessiva da cabeça, rotação exagerada do tronco e elevação dos ombros, acompanhadas de seu estado atual (ativado ou desativado).

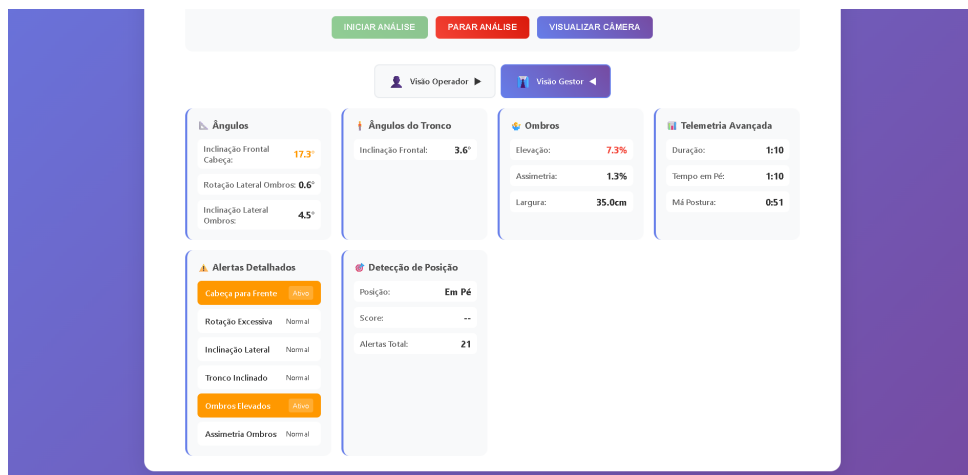


Figura 3.17: Modo de Análise em Tempo Real — Visão Gestor.

A interface também disponibiliza uma secção de Gestão de Usuários, ilustrada na Figura 3.18. Nela, é possível seleccionar e remover usuários previamente cadastrados. Essa ação é acompanhada de um aviso destacado, alertando que a exclusão é irreversível e resultará na eliminação completa dos dados de calibração e relatórios de telemetria associados. O utilizador conta ainda com o botão Atualizar Lista, que recarrega as informações armazenadas na base de dados, assegurando que as operações de gestão sejam realizadas sobre dados atualizados.

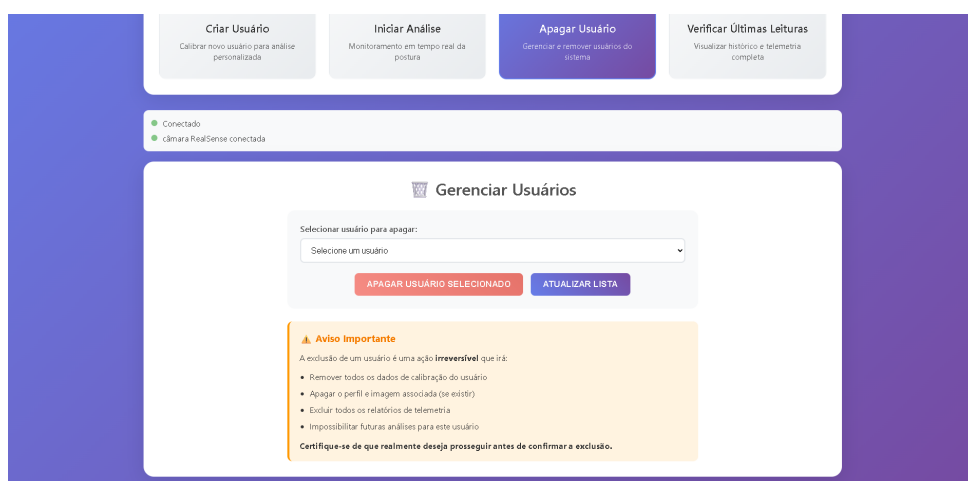


Figura 3.18: Tela de Gerenciamento e Exclusão de Usuários.

A sessão de Verificar últimas leituras contém as informações coletadas durante as sessões do sistema. Nela é possível verificar a média de tempo que o usuário ficou de pé ou sentado e a parcela de tempo que ficou em boa ou má postura. Também é mostrado quais eventos de má

postura foram ativados durante a sessão e as médias dos ângulos aferidos.

A Figura 3.19 representa a parte superior da página de análise de telemetria de sessão. Mostrando a seleção de usuário e sessão requerida. Também é possível ver o tempo da sessão, quantidade de alertas totais, tempo em má postura e tempo em pé.

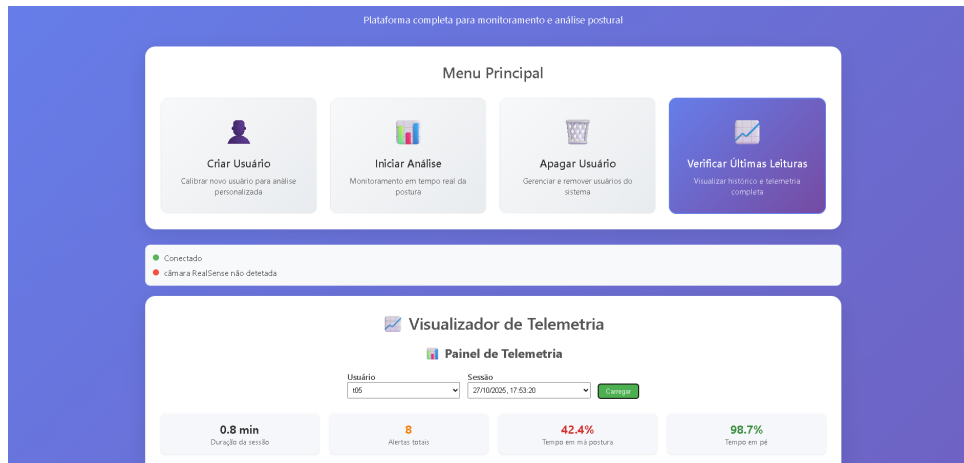


Figura 3.19: Tela de Visualização das Últimas Sessões Parte Superior.

A Figura 3.20 apresenta os gráficos de rosca, que são responsáveis por apresentar o Tempo em Pé e o tempo em boa postura. Abaixo dos gráficos de rosca tem dois gráficos de barra, representando a quantidade de eventos ativados durante a sessão e a média dos ângulos medidos.

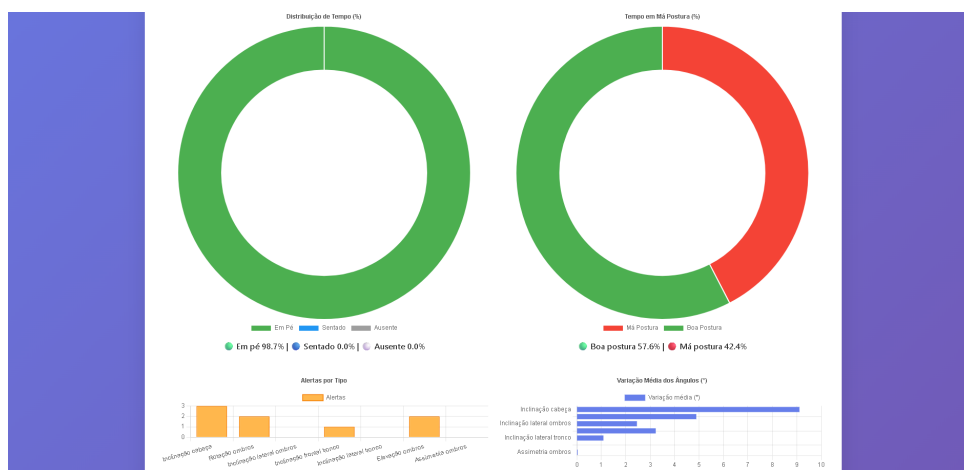


Figura 3.20: Tela de Visualização das Últimas Sessões Parte Inferior.

De modo geral, a interface gráfica constitui o elo de interação entre o usuário e o sistema de análise, traduzindo os cálculos biomecânicos e os dados de postura em informações visuais de

fácil interpretação. Além de simplificar o controlo das etapas de calibração e análise, ela garante operação totalmente local, respeitando as diretrizes de privacidade e segurança de dados descritas na sessão anterior.

O fluxo de comunicação entre o navegador, o servidor *Flask* e os módulos de análise está representado na Figura 3.21, através de um diagrama de sequência que ilustra a troca de eventos e dados em tempo real.

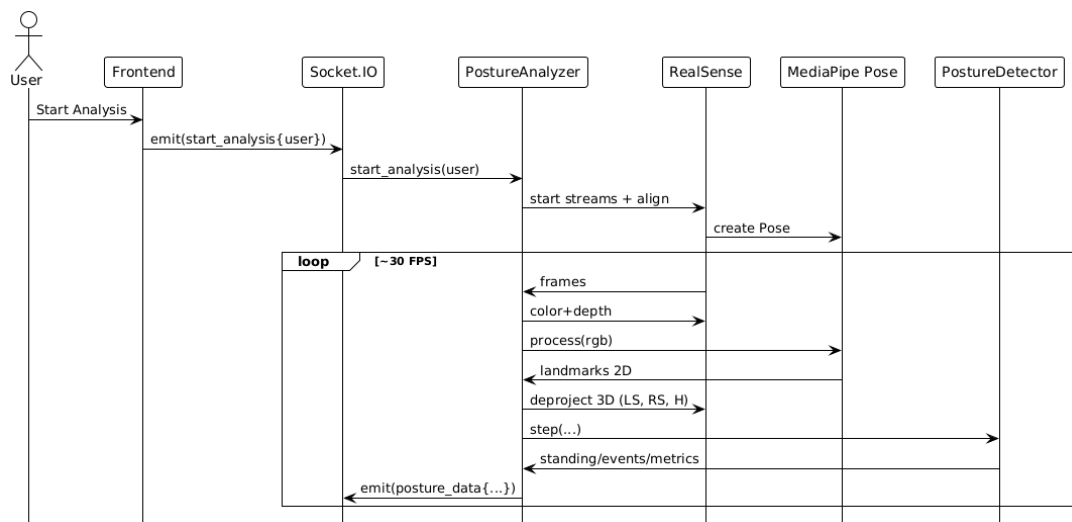


Figura 3.21: Diagrama de Comunicação da Camada *Web Server*.

3.4 Resumo

Esta sessão detalha a arquitetura e a estrutura do sistema, cujo fluxo de operação segue os seguintes passos:

O sistema de análise ergonómica foi desenvolvido com uma arquitetura modular que integra processamento em tempo real, filtragem de dados e uma interface web interativa. Seu funcionamento segue um pipeline lógico e contínuo, desde a criação do usuário até a geração dos alertas posturais, garantindo rastreabilidade, personalização e precisão nas medições.

O fluxo inicia-se no *frontend*, onde o operador cria um novo usuário e inicia o processo de calibração. Nesta etapa, o sistema gera automaticamente uma *baseline* individual, contendo os vetores de referência e parâmetros biomecânicos do usuário, que servem como ponto de comparação para todas as análises posteriores.

Com a calibração concluída, o sistema entra na fase de **análise em tempo real**, captando continuamente os pontos tridimensionais do corpo através da câmara e calculando os ângulos posturais relevantes. Os dados são então submetidos a um módulo de **filtragem e suavização**, composto por filtros de EMA e Mediana, que eliminam ruídos e estabilizam o sinal sem atrasar a resposta do sistema.

Em seguida, os valores filtrados são **comparados com a *baseline*** do usuário, identificando desvios posturais. Caso algum ângulo ultrapasse o limiar definido, entra em ação a **lógica de persistência temporal** (*Latch*), responsável por confirmar a consistência do desvio antes de gerar um evento. Se a persistência for validada, ocorre a **ativação do evento ergonômico**, que é imediatamente refletida na interface e registrada no banco de dados SQLite.

A **interface gráfica**, desenvolvida em ambiente web local (Flask + WebSocket), é o elo central de interação entre o usuário e o sistema. Ela oferece controle completo das etapas de calibração, análise e visualização dos resultados. O modo *Operador* apresenta dados simplificados e alertas visuais, enquanto o modo *Gestor* disponibiliza métricas detalhadas, gráficos de telemetria e histórico de sessões, facilitando o acompanhamento da performance ergonômica ao longo do tempo.

De forma geral, o sistema combina análise biomecânica em tempo real com visualização interativa e armazenamento estruturado de dados, formando um pipeline robusto e contínuo, da captação à decisão, capaz de apoiar o monitoramento ergonômico com precisão, estabilidade e fácil interpretação.

O código-fonte completo do sistema encontra-se disponível no Apêndice A.

Capítulo 4

Validação e Discussão

Neste capítulo será apresentada e discutido sobre a validação do sistema, uma comparação entre a saída do sistema com a avaliação de uma profissional da área da ergonomia no trabalho. Também será apresentado o desempenho em tempo real.

4.1 Validação do Sistema

A validação do sistema proposto foi conduzida mediante avaliação comparativa entre as classificações automáticas geradas pelo sistema e a avaliação realizada por especialista em ergonomia. O protocolo experimental consistiu na análise de 41 *frames* extraídos de sequência de vídeo, capturada durante a execução de atividades típicas de trabalho em bancada.

O estabelecimento do padrão de referência (*ground truth*) foi realizado por ergonomista com formação em Fisioterapia e 10 anos de experiência em análise postural ocupacional, aplicando sistematicamente os critérios do método RULA. A avaliação foi conduzida de forma independente, sem conhecimento prévio das classificações geradas pelo sistema proposto, assegurando a ausência de viés de confirmação.

4.1.1 Desempenho Geral do Sistema e Métricas Diagnósticas

Do total de 41 *frames* analisados, 37 (90,2%) apresentaram visual suficiente para avaliação pela especialista mediante análise da imagem RGB bidimensional. Em 4 *frames* (9,8%), a ergonomista

classificou a informação como insuficiente para determinação postural precisa, devido à limitação inerente à análise bidimensional para detecção de deslocamento no plano sagital.

A distribuição das posturas nos 37 *frames* avaliáveis foi de 29 casos classificados como inadequados (78,4%) e 8 casos classificados como adequados (21,6%) pela especialista, refletindo a natureza deliberada da seleção de *frames* para incluir representação equilibrada de ambas as categorias posturais.

A Tabela 4.1 apresenta a matriz de confusão obtida a partir da comparação entre as classificações do sistema e as avaliações da especialista.

Tabela 4.1: Matriz de confusão para detecção de posturas inadequadas ($n = 37$)

Sistema	Especialista		Total
	Adequada	Inadequada	
Adequada	8 (VN)	2 (FN)	10
Inadequada	0 (FP)	27 (VP)	27
Total	8	29	37

Nota: VP = Verdadeiro Positivo; VN = Verdadeiro Negativo; FP = Falso Positivo; FN = Falso Negativo.

Observa-se predominância de verdadeiros positivos (VP = 27) e ausência de falsos positivos, indicando que o sistema não classificou incorretamente nenhuma postura adequada como inadequada.

O sistema demonstrou acurácia global de 94,6% (35/37), com intervalo de confiança de 95% entre 81,8% e 99,3%, calculado pelo método de Wilson, recomendado para proporções em amostras pequenas. Este resultados indica elevada concordância entre as classificações automáticas e a avaliação especializada.

A Tabela 4.2 sumariza as métricas de desempenho diagnóstico do sistema, calculadas a partir da matriz de confusão apresentada.

Tabela 4.2: Métricas de desempenho do sistema de detecção ergonómica

Métrica	Fórmula	Valor	IC 95%	Interpretação
Sensibilidade (Recall)	$\frac{VP}{VP + FN}$	93,1%	[77,2–99,2%]	Excelente
Especificidade	$\frac{VN}{VN + FP}$	100%	[63,1–100%]	Perfeito
Acurácia	$\frac{VP + VN}{Total}$	94,6%	[81,8–99,3%]	Excelente
Precisão	$\frac{VP}{VP + FP}$	100%	[87,2–100%]	Perfeito
Valor Preditivo Negativo	$\frac{VN}{VN + FN}$	80,0%	[44,4–97,5%]	Bom
F1-Score	$2 \cdot \frac{Prec \cdot Sens}{Prec + Sens}$	96,4%	—	Excelente

Nota: IC = Intervalo de Confiança calculado pelo método de Wilson.

4.1.2 Análise Detalhada da Classificação e Falsos Negativos

O sistema apresentou especificidade de 100%, não registando nenhum falso positivo na amostra analisada. Este resultado demonstra que todas as classificações de postura inadequada geradas pelo sistema corresponderam efetivamente a posturas identificadas como ergonomicamente inadequadas pela especialista.

O valor preditivo positivo de 100% reforça a confiabilidade dos alertas gerados, indicando que, quando o sistema sinaliza uma postura inadequada, existe certeza absoluta quanto à presença de risco ergonómico real. Esta característica é particularmente relevante para a aplicabilidade prática do sistema, uma vez que minimiza a ocorrência de alarmes falsos que poderiam comprometer a aceitação da ferramenta em ambiente ocupacional.

A sensibilidade observada foi de 93,1% (27/29), indicando que o sistema detetou corretamente 27 das 29 posturas inadequadas identificadas pela especialista. Foram registados 2 falsos negativos, correspondendo a 5,4% da amostra total avaliável.

A investigação detalhada dos dois casos de falsos negativos revelou:

Caso FN-1: A ergonomista detetou uma “leve flexão cervical”, sendo representado pelo sistema como uma inclinação frontal da cabeça. Ao analisar os dados desta imagem, nota que

existe uma inclinação frontal de 4° , o que para a lógica de persistência e limiar do sistema desenvolvido para evitar pequenas flutuações, não detetou esta inclinação. Este falso negativo pode também ser interpretado como uma calibração de *baseline* imprecisa. O *frame* analisado é apresentado na Figura 4.1.



Figura 4.1: Caso 1 de Falso Negativo - Leve Flexão Cervical Detetado Pela Especialista. Sistema Retornou Boa Postura

Caso FN-2: Inadequação postural não detetada relacionada a rotação de cervical. A ergonomista classificou este *frame* como postura inadequada devido a uma rotação da coluna cervical (movimento de olhar para o lado), representado na Figura 4.2. Este *frame* não foi classificado pelo sistema pois o mesmo não deteta o movimento de rotação da cervical.

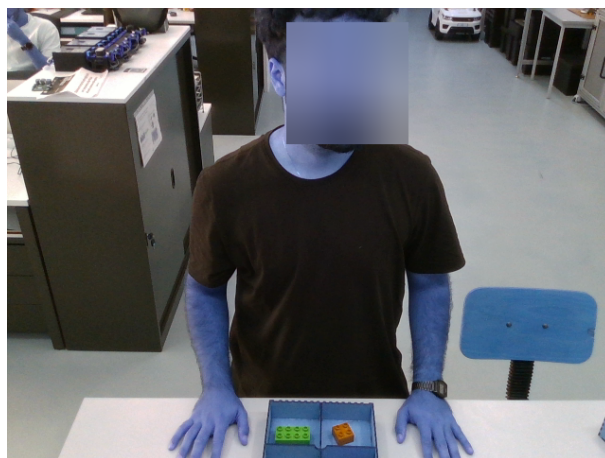


Figura 4.2: Caso 2 de Falso Negativo - Rotação de Cervical Detetado Pela Especialista. Sistema Retornou Boa Postura

4.1.3 Valor Diagnóstico da Informação de Profundidade

Nos 4 *frames* (9,8% da amostra total) classificados pela especialista como “informação insuficiente” para análise visual bidimensional, o sistema identificou posturas inadequadas em todos os casos, utilizando dados tridimensionais provenientes do sensor de profundidade.

A impossibilidade de avaliação visual nestes casos decorreu da limitação inerente à análise de imagens RGB frontais para detecção de deslocamentos angulares no plano sagital, particularmente flexão do tronco e inclinação frontal da cabeça. A capacidade do sistema de quantificar estes ângulos mediante informação de profundidade representa uma vantagem significativa em relação aos métodos tradicionais de avaliação ergonómica baseados em observação direta ou análise de vídeo bidimensional.

Este resultado evidencia que a incorporação de dados tridimensionais amplia substancialmente o espectro de posturas detetáveis, superando limitações metodológicas dos instrumentos convencionais de análise postural.

4.1.4 Síntese dos Resultados

Os resultados da validação demonstram desempenho robusto do sistema proposto, com acurácia superior a 94% e ausência de falsos positivos. A sensibilidade de 93,1% indica capacidade adequada de detecção de posturas inadequadas, embora a ocorrência de dois falsos negativos aponte para oportunidades de refinamento algorítmico ou ajuste de limiares.

A especificidade perfeita e o valor preditivo positivo de 100% constituem características especialmente relevantes para aplicações práticas, garantindo confiabilidade total dos alertas gerados e favorecendo a aceitação do sistema por parte dos usuários finais.

Adicionalmente, a capacidade de detecção de posturas inadequadas em situações onde a análise visual bidimensional é insuficiente demonstra valor agregado significativo da tecnologia de visão em profundidade para avaliação ergonómica automatizada.

4.2 Validação de Desempenho

A avaliação de desempenho do sistema foi conduzida medindo-se a taxa de processamento (FPS) e a latência média por *frame*. O FPS efetivo foi calculado a partir da razão entre o número total

de quadros processados e a duração da sessão de análise, conforme registrado na telemetria. A latência de processamento, por sua vez, foi estimada pela diferença temporal entre o início do ciclo de leitura da câmara e o envio dos resultados via *WebSocket*.

Os testes mostraram uma média de 13FPS e latência de 78ms por quadro, valores que garantem operação em tempo real para aplicações ergonômicas. Esses resultados confirmam que o pipeline de processamento, incluindo estimativa de pose, reconstrução 3D e cálculo de métricas, mantém desempenho estável e responsivo mesmo em ambiente local.

4.3 Discussão dos Resultados Obtidos

O sistema desenvolvido propõe uma solução não invasiva para a análise ergonômica em postos de trabalho em pé, executando-se integralmente na CPU, o que amplia sua portabilidade e acessibilidade, possibilitando a operação em equipamentos sem GPU dedicada para o processamento de imagens.

Diante do aumento dos casos de LMERT, a solução proposta demonstra potencial como ferramenta de apoio a profissionais da saúde ocupacional e ergonomia, permitindo o monitoramento contínuo da postura dos trabalhadores durante suas atividades. O sistema gera resumos detalhados e mantém uma base de dados estruturada, possibilitando análises longitudinais e a detecção precoce de padrões de risco postural.

A presença de uma interface interativa e de alertas automáticos em tempo real amplia sua aplicabilidade prática, configurando-o como um recurso complementar às avaliações presenciais conduzidas por ergonomistas e fisioterapeutas ocupacionais. Assim, o sistema contribui para uma abordagem proativa e preventiva na gestão da saúde ocupacional, favorecendo a redução de afastamentos laborais e a melhoria da qualidade de vida no trabalho.

Os resultados obtidos indicam que o sistema pode ser integrado a ambientes industriais inteligentes, alinhando-se às diretrizes da Indústria 5.0, que prioriza o bem-estar humano e a sustentabilidade.

O desempenho do sistema demonstrou elevada dependência da calibração da postura neutra, que deve ser executada de forma precisa e em condições equivalentes ao posto de trabalho real. Uma calibração imprecisa ou realizada em ambiente distinto pode comprometer a acurácia da referência postural, afetando diretamente a confiabilidade das medições subsequentes.

Esse procedimento pode demandar assistência externa para garantir o correto posicionamento e enquadramento do operador durante a etapa de calibração.

Além disso, o ângulo de posicionamento da câmara constitui uma limitação relevante. A configuração atual ocasiona oclusão parcial da região dos quadris, que varia conforme a estatura do usuário, e oclusão total dos membros inferiores. Essas restrições visuais reduzem a capacidade de análise postural completa e podem, mesmo com o mecanismo de correção de quadris implementado, introduzir erros residuais de medição em determinadas condições de operação.

4.4 Considerações Éticas e Legais sobre o Uso de Imagens e Dados Pessoais

Durante o desenvolvimento do sistema, foram adotadas medidas rigorosas para assegurar a conformidade com princípios éticos e legais relacionados à privacidade e à proteção de dados pessoais.

O sistema não realiza qualquer forma de armazenamento ou transmissão de imagens que permitam a identificação do usuário. As imagens captadas pela câmara Intel RealSense D415 são utilizadas exclusivamente em memória volátil (*RAM*) durante o processamento e são descartadas imediatamente após a extração dos pontos corporais tridimensionais. Nenhum dado visual, como fotografias, vídeos ou traços faciais, é gravado em disco, armazenado em banco de dados ou transmitido pela rede.

O modelo de detecção de pose (*MediaPipe Pose*) atua apenas sobre as coordenadas das *landmarks* corporais, e não sobre características faciais ou biométricas que possam identificar o indivíduo. Os dados armazenados em banco de dados (*SQLite*) correspondem exclusivamente a valores numéricos anonimizados, como ângulos articulares, índices de postura e parâmetros derivados da calibração.

Dessa forma, o sistema garante que nenhuma informação pessoal sensível seja coletada, tratada ou divulgada, em conformidade com a legislação europeia de proteção de dados, o General Data Protection Regulation — Regulamento Geral sobre a Proteção de Dados (GDPR) [48], e com a legislação brasileira correspondente, a Lei Geral de Proteção de Dados Pessoais — Lei nº 13.709, de 14 de agosto de 2018 (LGPD) [49]. Ambas as normas são baseadas em

princípios de transparência, consentimento informado, minimização e anonimização de dados pessoais.

Por fim, todas as análises são realizadas de forma local, sem necessidade de conexão à *internet*, e o objetivo do sistema é estritamente ergonómico, não envolvendo qualquer tipo de reconhecimento, monitorização ou identificação pessoal.

4.5 Resumo

Este capítulo apresentou o processo de validação e a discussão dos resultados do sistema proposto. A validação, realizada através da comparação com a avaliação de uma ergonomista especialista, demonstrou um desempenho robusto, com acurácia global de 94,6%. Destaca-se a especificidade e precisão de 100%, indicando que o sistema não gerou falsos positivos, um fator crítico para a confiabilidade e aceitação da ferramenta em ambiente ocupacional. A sensibilidade de 93,1% foi considerada excelente, e a análise dos dois falsos negativos registados apontou para limitações conhecidas (movimentos não monitorizados, como a rotação cervical) e oportunidades de refinamento nos limiares de deteção.

Um dos resultados mais significativos foi a capacidade do sistema em detetar posturas inadequadas em *frames* onde a análise 2D pela especialista foi considerada "insuficiente", comprovando o valor diagnóstico do sensor de profundidade para movimentos no plano sagital. O desempenho em tempo real foi validado, com o sistema a operar a uma média de 13 FPS em equipamentos sem GPU dedicada. A discussão final reforçou o potencial do sistema como uma ferramenta acessível (baseada em CPU) e alinhada à Indústria 5.0, mas também sublinhou as suas limitações, como a alta dependência de uma calibração precisa e as oclusões parciais impostas pelo ângulo da câmara.

Capítulo 5

Conclusão

A transição para a Indústria 5.0 estabeleceu o bem-estar humano como um pilar central dos processos produtivos. Contudo, a prevalência das LMERTs permanece como um dos principais desafios à saúde ocupacional, gerando custos sociais e económicos significativos. A origem destas lesões está frequentemente ligada a fatores de risco ergonómico, como a manutenção de posturas inadequadas por períodos prolongados. Historicamente, a avaliação destes riscos dependia de métodos de observação subjetivos e pontuais ou de sensores “vestíveis” invasivos, que podem alterar a própria postura que se pretende medir. Este cenário criou a necessidade de uma nova abordagem, explorada nesta dissertação: o desenvolvimento de um sistema não-invasivo, baseado em visão computacional, capaz de monitorizar a postura em tempo real e de forma preventiva.

5.1 Considerações Finais

O sistema desenvolvido nesta dissertação atendeu aos objetivos inicialmente propostos, integrando de forma coerente todos os módulos planeados. O módulo de deteção de pose foi implementado com base em técnicas de visão computacional, possibilitando a extração de ângulos posturais a partir de imagens em tempo real. A interface de apresentação desenvolvida mostrou-se funcional e intuitiva, permitindo o acompanhamento visual imediato dos resultados, bem como a emissão de alertas automáticos em situações de risco postural.

O projeto atingiu ainda o objetivo de construir um sistema não invasivo de análise ergonómica, capaz de operar exclusivamente em CPU, o que amplia sua portabilidade e viabiliza o uso

em diferentes contextos industriais e laboratoriais. A validação com especialista demonstrou elevada concordância entre o sistema e a avaliação humana, confirmando a eficácia do método proposto para aplicação prática.

De forma mais ampla, esta dissertação contribui para o avanço do paradigma da Indústria 5.0, ao incorporar o bem-estar humano como elemento central dos processos produtivos. Ao monitorizar e avaliar continuamente a postura dos trabalhadores, o sistema proposto promove a prevenção de LMERT, reforçando o papel da tecnologia como ferramenta de apoio à saúde ocupacional e à sustentabilidade do ambiente de trabalho.

5.2 Trabalhos Futuros

Como continuidade deste trabalho, recomenda-se o desenvolvimento de aprimoramentos que ampliem a precisão e o escopo de aplicação do sistema proposto. Entre as possibilidades, destaca-se a inclusão de novos ângulos de medição, como a rotação da cervical e o ângulo dos membros superiores, o que permitiria uma análise postural mais abrangente e detalhada, especialmente em atividades que envolvem movimentos repetitivos dos braços e pescoço.

Outra vertente relevante consiste em padronizar o processo de calibração, de modo a eliminar a necessidade de definição manual da postura neutra pelo usuário. Tal aprimoramento poderia ser alcançado por meio de modelos estatísticos ou aprendizado de máquina, capazes de identificar automaticamente a postura de referência com base em padrões corporais médios. Essa automatização aumentaria a usabilidade e reduziria erros associados à calibração manual.

Outra direção importante refere-se ao aprofundamento da fase de calibração e da validação experimental. Recomenda-se a realização de testes com uma amostra maior e mais diversificada de utilizadores, contemplando diferentes biotipos, estaturas e condições reais de operação. Além disso, é pertinente conduzir validações presenciais com um profissional de ergonomia acompanhando o sistema em funcionamento, possibilitando a comparação direta, em tempo real, entre a avaliação humana e a saída automatizada. Essa abordagem permitiria identificar eventuais discrepâncias, ajustar limiares e refinar os modelos utilizados. Adicionalmente, a definição de protocolos padronizados para a coleta dos dados e execução das tarefas contribuiria para maior rigor metodológico e reprodutibilidade dos resultados.

Além disso, sugere-se a implementação de um módulo para cálculo de repetitividade dos

movimentos, permitindo avaliar a frequência e a duração das posturas adotadas ao longo da jornada de trabalho. Essa funcionalidade contribuiria para a avaliação de risco cumulativo, complementando as medições angulares com indicadores de exposição prolongada a posturas inadequadas.

Esses aprimoramentos futuros têm o potencial de tornar o sistema mais robusto, autônomo e abrangente, fortalecendo seu papel como ferramenta de apoio à ergonomia e à prevenção de LMERT.

Bibliografia

- [1] E. Commission, D.-G. for Research, Innovation, M. Breque, L. De Nul e A. Petridis, *Industry 5.0 – Towards a sustainable, human-centric and resilient European industry*. Publications Office of the European Union, 2021. DOI: doi/10.2777/308407.
- [2] Y. Lu et al., “Outlook on human-centric manufacturing towards Industry 5.0,” *Journal of Manufacturing Systems*, vol. 62, pp. 612–627, fev. de 2022. DOI: 10.1016/j.jmsy.2022.02.001.
- [3] A. H. Maslow, “A theory of human motivation,” *Psychological Review*, vol. 50, n.º 4, pp. 370–396, 1943. DOI: 10.1037/h0054346.
- [4] European Agency for Safety and Health at Work. “Musculoskeletal disorders.” European Agency for Safety and Health at Work website, acessido em 6 de out. de 2025. URL: <https://osha.europa.eu/pt/themes/musculoskeletal-disorders>.
- [5] UGT, *Lesões Músculo-Esqueléticas relacionadas com o Trabalho: Riscos, Efeitos na Saúde e Prevenção*, Guia temático, Acesso em: 3 out. 2025, nov. de 2017. URL: <https://www.ugt.pt/publicfiles/kk1geivzmvhbbqvxoakdagb9jndzlhss61fj5kfr.pdf>.
- [6] European Agency for Safety and Health at Work, “Work-related Musculoskeletal Disorders: Prevalence, Costs and Demographics in the EU,” European Agency for Safety e Health at Work (EU-OSHA), Luxembourg, 2019, Publications Office of the European Union. acessido em 6 de out. de 2025. URL: https://osha.europa.eu/sites/default/files/Work-related_MSDs_prevalence_costs_and_demographics_in_the_EU_report.pdf.

- [7] BR MED, *LER/DORT: Como lesões ocupacionais comprometem a produtividade no Brasil*, Blog BR MED, Publicado em 11 de novembro de 2024. Acesso em: 7 out. 2025, nov. de 2024. URL: <https://grupobrmed.com.br/blog/ler-dort-afetam-produtividade-de-profissionais-no-brasil/>.
- [8] G.-Y. Kim et al., “Human Digital Twin System for Operator Safety and Work Management,” em set. de 2022, pp. 529–536, ISBN: 978-3-031-16410-1. DOI: 10.1007/978-3-031-16411-8_61.
- [9] C.-T. Kuo et al., “Human Posture Transition-Time Detection Based upon Inertial Measurement Unit and Long Short-Term Memory Neural Networks,” *Biomimetics*, vol. 8, n.º 6, p. 471, 2023, ISSN: 2313-7673. DOI: 10.3390/biomimetics8060471. URL: <https://www.mdpi.com/2313-7673/8/6/471>.
- [10] T. Agostinelli, A. Generosi, S. Ceccacci, R. Menghi, M. Germani e A. Monteriù, “Validation of computer vision-based ergonomic risk assessment tools for real manufacturing environments,” *Scientific Reports*, vol. 14, p. 27785, 2024. DOI: 10.1038/s41598-024-79373-4. URL: <https://doi.org/10.1038/s41598-024-79373-4>.
- [11] S. Nahavandi, “Industry 5.0—A Human-Centric Solution,” *Sustainability*, vol. 11, n.º 16, p. 4371, 2019. DOI: 10.3390/su11164371. URL: <https://doi.org/10.3390/su11164371>.
- [12] P. K. R. Maddikunta et al., “Industry 5.0: A survey on enabling technologies and potential applications,” *Journal of Industrial Information Integration*, vol. 26, p. 100257, 2022. DOI: 10.1016/j.jii.2021.100257. URL: <https://doi.org/10.1016/j.jii.2021.100257>.
- [13] International Ergonomics Association, *What is Ergonomics*, <https://iea.cc/about/what-is-ergonomics/>, Acedido em: 10 de outubro de 2025, 2024.
- [14] World Health Organization, *Musculoskeletal conditions*, <https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions>, Acedido em: 10 de outubro de 2025, 2019.

- [15] B. R. da Costa e E. R. Vieira, “Risk factors for work-related musculoskeletal disorders: A systematic review of the literature,” *American Journal of Industrial Medicine*, vol. 53, n.º 3, pp. 285–323, 2010. DOI: 10.1002/ajim.20750. URL: <https://doi.org/10.1002/ajim.20750>.
- [16] A. A. Shikdar e M. A. Al-Kindi, “Office Ergonomics: Deficiencies in Computer Workstation Design,” *International Journal of Occupational Safety and Ergonomics*, vol. 13, n.º 2, pp. 215–223, 2007. DOI: 10.1080/10803548.2007.11076722. URL: <https://doi.org/10.1080/10803548.2007.11076722>.
- [17] P. T. Katzmarzyk et al., “Sedentary Behavior and Health: Update from the 2018 Physical Activity Guidelines Advisory Committee,” *Medicine & Science in Sports & Exercise*, vol. 51, n.º 6, p. 1227, 2019.
- [18] European Agency for Safety and Health at Work (EU-OSHA), *Work-related musculoskeletal disorders: facts and figures*, <https://osha.europa.eu/en/publications/work-related-musculoskeletal-disorders-facts-and-figures>, ISBN: 978-92-9479-223-1, 2019.
- [19] Ministério da Saúde do Brasil, *Protocolo de Dor Relacionada ao Trabalho: Lesões por Esforços Repetitivos (LER) e Distúrbios Osteomusculares Relacionados ao Trabalho (Dort)*, Série A. Normas e Manuais Técnicos. ISBN: 978-85-334-2615-5, Brasília, DF, 2018.
- [20] F. P. Kendall, E. K. McCreary, P. G. Provance, M. M. Rodgers e W. A. Romani, *Músculos: Provas e Funções, com Postura e Dor*, 5ª. São Paulo: Editora Manole, 2007, ISBN: 978-85-204-2432-2.
- [21] L. McAtamney e E. N. Corlett, “RULA: a survey method for the investigation of work-related upper limb disorders,” *Applied Ergonomics*, vol. 24, n.º 2, pp. 91–99, 1993, ISSN: 0003-6870. DOI: 10.1016/0003-6870(93)90080-S.

- [22] S. Hignett e L. McAtamney, “Rapid Entire Body Assessment (REBA),” *Applied Ergonomics*, vol. 31, n.º 2, pp. 201–205, 2000, ISSN: 0003-6870. DOI: 10.1016/S0003-6870(99)00039-3.
- [23] O. Karhu, P. Kansi e I. Kuorinka, “Correcting working postures in industry: A practical method for analysis,” *Applied Ergonomics*, vol. 8, n.º 4, pp. 199–201, 1977, ISSN: 0003-6870. DOI: 10.1016/0003-6870(77)90164-8.
- [24] Y.-J. Kwon, D.-H. Kim, B.-C. Son, K.-H. Choi, S. Kwak e T. Kim, “A Work-Related Musculoskeletal Disorders (WMSDs) Risk-Assessment System Using a Single-View Pose Estimation Model,” *International Journal of Environmental Research and Public Health*, vol. 19, n.º 16, p. 9803, 2022, Acesso em: 14 out. 2025, ISSN: 1660-4601. DOI: 10.3390/ijerph19169803. URL: <https://www.mdpi.com/1660-4601/19/16/9803>.
- [25] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer Science & Business Media, 2010, ISBN: 978-1848829343.
- [26] Microsoft, *Azure Kinect DK*, <https://azure.microsoft.com/en-us/services/kinect-dk/>, Acedido em: 19 de outubro de 2025, out. de 2025.
- [27] Intel, *Intel RealSense*, <https://www.intelrealsense.com/>, Acedido em: 19 de outubro de 2025, out. de 2025.
- [28] M. T. Talacio, “Development of an Intelligent Personal Assistant to Empower Operators in Industry 4.0 Environments,” Acedido em: 10 de outubro de 2025, tese de mestrado, Instituto Politécnico de Bragança, Bragança, Portugal, 2020. URL: <http://hdl.handle.net/10198/23238>.
- [29] S. Grigorescu, B. Trasnea, T. Cocias e G. Macesanu, “A survey of deep learning techniques for autonomous driving,” *Journal of Field Robotics*, vol. 37, n.º 3, pp. 362–386, 2020. DOI: 10.1002/rob.21918. URL: <https://doi.org/10.1002/rob.21918>.

- [30] Y. Taigman, M. Yang, M. Ranzato e L. Wolf, “DeepFace: Closing the Gap to Human-Level Performance in Face Verification,” em *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE Computer Society, 2014, pp. 1701–1708. DOI: 10.1109/CVPR.2014.220. URL: <https://doi.org/10.1109/CVPR.2014.220>.
- [31] T. Czimmermann et al., “Visual-Based Defect Detection and Classification Approaches for Industrial Applications — A SURVEY,” *Sensors*, vol. 20, n.º 5, p. 1459, 2020. DOI: 10.3390/s20051459. URL: <https://doi.org/10.3390/s20051459>.
- [32] A. Kamilaris e F. X. Prenafeta-Boldú, “Deep Learning in Agriculture: A Survey,” *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018. DOI: 10.1016/j.compag.2018.02.016. URL: <https://doi.org/10.1016/j.compag.2018.02.016>.
- [33] G. Litjens et al., “A Survey on Deep Learning in Medical Image Analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017. DOI: 10.1016/j.media.2017.07.005. URL: <https://doi.org/10.1016/j.media.2017.07.005>.
- [34] Y. Chen, Z. Wang, Y. Peng, Z. Zhang, G. Yu e J. Sun, “A survey on 2D human pose estimation with deep learning,” *ACM Computing Surveys (CSUR)*, vol. 53, n.º 2, pp. 1–37, 2020, ISSN: 0360-0300. DOI: 10.1145/3380463.
- [35] Z. Cao, T. Simon, S.-E. Wei e Y. Sheikh, “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” em *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2017, pp. 1302–1310. DOI: 10.1109/CVPR.2017.143. URL: <https://doi.org/10.1109/CVPR.2017.143>.
- [36] Google, *Pose Landmarker - MediaPipe*, https://ai.google.dev/edge/mediapipe/solutions/vision/pose_landmarker?hl=pt-br, Acedido em: 19 de outubro de 2025, out. de 2025.
- [37] V. Bazarevsky, I. Grishchenko et al., “BlazePose: On-device Real-time Body Pose tracking,” *arXiv preprint arXiv:2006.10204*, 2020.

- [38] Ultralytics, *Pose Estimation - Ultralytics Docs*, <https://docs.ultralytics.com/pt/tasks/pose/>, Acedido em: 19 de outubro de 2025, out. de 2025.
- [39] G. Jocher et al., *Ultralytics YOLOv8*, <https://github.com/ultralytics/ultralytics>, 2023.
- [40] M. Corporation, *Kinect for Xbox 360 Launch Announcement*, <https://news.microsoft.com/2010/06/13/kinect-for-xbox-360-launches/>, Acessado em 2025-02-10, 2010.
- [41] V. M. Manghisi, A. E. Uva, M. Fiorentino, V. Bevilacqua, G. F. Trotta e G. Monno, “Real time RULA assessment using Kinect v2 sensor.,” *Applied ergonomics*, vol. 65, pp. 481–491, 2017. URL: <https://api.semanticscholar.org/CorpusID:42119357>.
- [42] H. Haggag, M. Hossny, S. Nahavandi e D. C. Creighton, “Real Time Ergonomic Assessment for Assembly Operations Using Kinect,” *2013 UKSim 15th International Conference on Computer Modelling and Simulation*, pp. 495–500, 2013. URL: <https://api.semanticscholar.org/CorpusID:16055156>.
- [43] L. Li, T. Martin e X. Xu, “A novel vision-based real-time method for evaluating postural risk factors associated with musculoskeletal disorders,” *Applied Ergonomics*, vol. 87, p. 103 138, 2020, Acesso em: 14 out. 2025, ISSN: 0003-6870. DOI: 10.1016/j.apergo.2020.103138. URL: <https://www.sciencedirect.com/science/article/pii/S0003687020300958>.
- [44] M. Varas, L. Chang e J. C. Garcia, “Risk Assessment of Musculoskeletal Disorders Using Artificial Intelligence,” em *E3S Web of Conferences, Volume 532 — CIIA 2024*, Acesso em: 14 out. 2025, 2024, p. 02001. DOI: 10.1051/e3sconf/202453202001. URL: https://www.e3s-conferences.org/articles/e3sconf/pdf/2024/62/e3sconf_ciia2024_02001.pdf.

- [45] LearnOpenCV. “YOLOv7-Pose vs MediaPipe in Human Pose Estimation.” Acesso em: 14 out. 2025, LearnOpenCV. URL: <https://learnopencv.com/yolov7-pose-vs-mediapipe-in-human-pose-estimation/>.
- [46] SEPRI, *Postura Forçada Em Pé Por Períodos Prolongados: o que é, quais os efeitos e como prevenir*, Acesso em: 14 out. 2025, jul. de 2022. URL: <https://sepri.pt/noticia/postura-forcada-em-pe-por-periodos-prolongados/>.
- [47] European Agency for Safety and Health at Work (EU-OSHA), *Musculoskeletal disorders and prolonged static standing*, Acesso em: 19 de Outubro de 2025, mai. de 2020. URL: <https://oshwiki.osha.europa.eu/en/themes/musculoskeletal-disorders-and-prolonged-static-standing>.
- [48] E. Union, *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)*, Official Journal of the European Union, L 119, 4 May 2016, Accessed: 2025-10-24, 2016. URL: <https://eur-lex.europa.eu/eli/reg/2016/679>.
- [49] Brasil, *Lei nº 13.709, de 14 de agosto de 2018 — Lei Geral de Proteção de Dados Pessoais*, Diário Oficial da União, Brasília, DF, 15 ago. 2018, Acesso em: 2025-10-24, 2018. URL: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm.

Apêndice A

Código Fonte


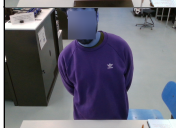


O código-fonte completo do sistema encontra-se disponível no repositório público:

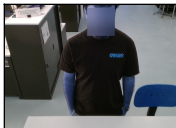
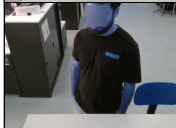











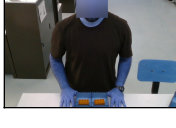
<https://github.com/pedropcsantos/posture-analysis-system>.

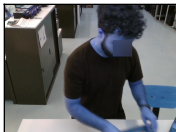



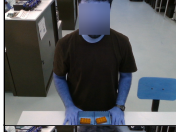








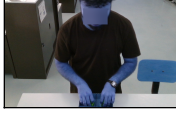
Apêndice B

Validação com Ergonomista

Este apêndice apresenta o relatório completo de validação do sistema desenvolvido, contendo todas as páginas do documento original gerado junto a ergonomista.

Imagem Ref.	Saída do Sistema	Avaliação Ergonomista	Comentários da Ergonomista	Resultado
	Adequada	Adequada	--	Concordância
	Inadequada	Inadequada	--	Concordância
	Inadequada	Inadequada	--	Concordância
	Inadequada	Inadequada	--	Concordância
	Inadequada	Inadequada	A assimetria de ombros pode se associada ou não a inclinação lateral do tronco com uma possível escoliose, e ou esta inclinação pode ser somente uma tensão muscular unilateral onde não ocorre a inclinação lateral do tronco.	Concordância
	Inadequada	Inadequada	idem frame 6	Concordância
	Inadequada	Inadequada	--	Concordância
	Inadequada	Inadequada	este frame aparenta uma leve assimetria de ombros	Concordância
	Adequada	Adequada	--	Concordância
	Adequada	Adequada	--	Concordância
	Inadequada	Inadequada	Leve flexão de coluna cervical	Concordância
	Inadequada	Inadequada	Rotação lateral de tronco	Concordância
	Inadequada	Inadequada	--	Concordância

	Inadequada	Insuficiente	--	Divergência
	Inadequada	Inadequada	Rotação lateral de tronco	Concordância
	Adequada	Inadequada	Leve flexão de coluna cervical	Falso Negativo (Erro do Sistema)
	Adequada	Adequada	--	Concordância
	Inadequada	Inadequada	flexão de coluna cervical e coluna dorsal	Concordância
	Inadequada	Inadequada	Leve flexão de coluna cervical com moderada de coluna lombar	Concordância
	Inadequada	Insuficiente	--	Divergência
	Inadequada	Inadequada	Leve flexão de coluna cervical com moderada de coluna lombar	Concordância
	Adequada	Adequada	--	Concordância
	Inadequada	Inadequada	Flexão com inclinação de tronco, flexão com abdução de ombro	Concordância
	Adequada	Adequada	--	Concordância
	Inadequada	Inadequada	Flexão de coluna cervical	Concordância
	Inadequada	Inadequada	Flexão com rotação de tronco e flexão de ombro bilateral	Concordância
	Adequada	Adequada	--	Concordância

	Inadequada	Inadequada	Rotação de tronco com leve flexão de coluna cervical	Concordância
	Inadequada	Inadequada	Rotação com flexão e inclinação de tronco	Concordância
	Inadequada	Insuficiente	--	Divergência
	Inadequada	Inadequada	Inclinação lateral de tronco com flexão e abdução de ombro	Concordância
	Adequada	Adequada	--	Concordância
	Inadequada	Inadequada	Rotação de tronco	Concordância
	Inadequada	Inadequada	Rotação com flexão e inclinação de tronco	Concordância
	Inadequada	Inadequada	Flexão de coluna cervical	Concordância
	Adequada	Inadequada	Rotação de cervical	Falso Negativo (Erro do Sistema)
	Inadequada	Inadequada	Flexão, inclinação de tronco com flexão e abdução de ombro	Concordância
	Inadequada	Inadequada	Flexão de tronco com flexão de ombro	Concordância
	Inadequada	Inadequada	Flexão de tronco com flexão de ombro	Concordância
	Inadequada	Inadequada	--	Concordância
	Inadequada	Insuficiente	--	Divergência