

ANÁLISE BIOMECÂNICA DE EXERCÍCIOS FÍSICOS EM TEMPO REAL: DETECÇÃO DE DESVIOS POSTURAIS COM MÉTRICAS BIOMECÂNICAS EM FLUXO CONTÍNUO

Deivid Crepaldi Campos¹, Marlon de Matos de Oliveira²

Resumo: A prática de exercícios físicos sem supervisão presencial tem crescido de forma expressiva, aumentando o risco de lesões por execução inadequada. A visão computacional surge como alternativa acessível para monitorar a biomecânica do movimento, e estudos prévios já exploraram essa via: Hussain et al. (2024) classificaram exercícios fisioterapêuticos com 99,6% de acurácia, Ko, Nasridinov e Park (2024) corrigiram a postura do agachamento em *powerlifting* com acurácia superior a 98%, e Rodriguez e Oliveira (2025) aplicaram MediaPipe e LSTM a poucos exercícios, sem entregar aplicação funcional em tempo real. Esses trabalhos concentram-se em exercícios dinâmicos isolados ou contextos específicos, deixando uma lacuna quanto a exercícios isométricos estáticos e a sistemas completos de captura, inferência e *feedback* em tempo real. Diante disso, este trabalho desenvolveu um sistema para análise biomecânica de exercícios de calistenia em tempo real, capaz de detectar desvios posturais e fornecer *feedback* corretivo sem sensores especializados. Desenvolvido em Python, com MediaPipe para estimativa de pose e redes LSTM para classificação temporal, o sistema foi treinado em 342 vídeos de quatro exercícios, com arquitetura em cascata (detector global e avaliadores por exercício). Os modelos atingiram acurácia entre 83,1% e 100% no conjunto de teste, e a validação funcional resultou em 87,8% de acurácia geral, contribuindo com exercícios isométricos pouco explorados na literatura.

Palavras-chave: Visão computacional; Estimativa de pose humana; Calistenia; Redes LSTM; Análise biomecânica; *Feedback* em tempo real.

ABSTRACT: Abstract: The unsupervised practice of physical exercises has grown significantly, increasing the risk of injuries due to improper execution. Computer vision emerges as an accessible alternative to monitor

¹Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense (Unesc), Criciúma - Santa Catarina - Brasil. deividccampos3@unesc.net

²Orientador, Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense (Unesc), Criciúma - Santa Catarina - Brasil. marlon.oliveira@unesc.net

movement biomechanics, and prior studies have already explored this approach: Hussain et al. (2024) classified physiotherapy exercises with 99.6% accuracy, Ko, Nasridinov e Park (2024) corrected squat posture in powerlifting with accuracy above 98%, and Rodriguez e Oliveira (2025) applied MediaPipe and LSTM to a small set of exercises, without delivering a functional real-time application. These works focus on isolated dynamic exercises or specific contexts, leaving a gap regarding static isometric exercises and complete systems integrating capture, inference, and real-time feedback. Given this, the present work developed a system for real-time biomechanical analysis of calisthenics exercises, capable of detecting postural deviations and providing corrective feedback without specialized sensors. Developed in Python, with MediaPipe for pose estimation and LSTM networks for temporal classification, the system was trained on 342 videos of four exercises, using a cascade architecture (global detector and per-exercise evaluators). The models achieved accuracy between 83.1% and 100% on the test set, and functional validation resulted in an overall accuracy of 87.8%, contributing static isometric exercises still underexplored in the literature.

Keywords: Computer vision; Human pose estimation; Calisthenics; LSTM networks; Biomechanical analysis; Real-time feedback.

1 INTRODUÇÃO

A evolução das tecnologias digitais tem proporcionado avanços significativos em diversas áreas, e a ciência da computação se destaca como eixo central desse processo. A visão computacional tem ganhado destaque ao possibilitar a interpretação de imagens e vídeos de forma automatizada, simulando aspectos da visão humana com precisão crescente e em tempo real. Esse avanço tem permitido aplicações em saúde, bem-estar e reabilitação, viabilizando o acompanhamento da execução de exercícios físicos em ambientes domésticos com o uso de câmeras comuns, como webcams e smartphones, democratizando o acesso e reduzindo custos (Cao et al., 2017).

De forma paralela, observa-se um fenômeno social relevante: o aumento expressivo do número de pessoas que realizam exercícios físicos sem supervisão presencial, motivadas por vídeos, aplicativos e programas de treinamento genéricos. Embora a prática regular de atividade física seja amplamente reconhecida pelos benefícios à saúde pública, a ausência de orientação pode resultar em execução inadequada, sobrecargas e

aumento do risco de lesões musculoesqueléticas, especialmente entre iniciantes (World Health Organization, 2020).

Estudos reforçam essa preocupação. Pesquisas apontam que técnica inadequada, progressão acelerada e cargas mal ajustadas estão entre os principais fatores associados ao surgimento de lesões em academias de musculação e em modalidades de alta intensidade. Ombros, região lombar e joelhos são frequentemente afetados, ressaltando a importância de mecanismos de *feedback* sobre a forma de execução (Bonilla et al., 2022; Szajkowski et al., 2023).

Nesse contexto, os exercícios com o peso corporal, conhecidos como calistenia, têm se destacado como uma abordagem eficiente e acessível para o desenvolvimento de força, resistência e controle motor. Essa modalidade utiliza a própria massa corporal como resistência, promovendo o recrutamento coordenado de múltiplos grupos musculares e a melhora da estabilidade articular (Iversen; Bø; Fimland, 2023). Por dispensarem o uso de equipamentos, os exercícios calistênicos possibilitam treinos em diferentes ambientes, favorecendo a adesão e a continuidade da prática física, aspecto relevante diante do aumento de treinamentos realizados em casa ou com suporte digital.

Além da praticidade, os exercícios com o peso corporal demandam elevada participação do sistema neuromuscular, uma vez que o controle do centro de massa e a estabilização segmentar dependem diretamente da coordenação intermuscular. Evidências mostram que treinos progressivos exclusivamente com peso corporal promovem adaptações significativas na força e na arquitetura muscular, com resultados comparáveis aos obtidos em programas com cargas externas (Wei et al., 2023). Esses achados reforçam que a calistenia constitui um método eficiente para o desenvolvimento físico, explorando alavancas corporais e manipulação do próprio peso para modular a intensidade.

Do ponto de vista biomecânico, entretanto, a execução desses exercícios apresenta grande variabilidade técnica. Movimentos como agachamentos, flexões ou avanços exigem coordenação adequada entre segmentos corporais e alinhamento postural preciso. Variações na largura de base, profundidade ou amplitude podem alterar significativamente os ângulos articulares e a distribuição de cargas no corpo (Mause et al., 2024). Assim, a qualidade da execução torna-se um fator determinante tanto para o desempenho quanto para a prevenção de lesões.

Nesse cenário, a estimativa de pose humana (HPE) consolidou-

se como uma tecnologia promissora para a análise automatizada de movimentos corporais. Modelos modernos de HPE, como OpenPose, BlazePose e MoveNet, permitem detectar pontos anatômicos do corpo humano a partir de câmeras comuns, possibilitando o cálculo automático de ângulos articulares, amplitude de movimento e padrões de execução (Lan et al., 2023). Essa abordagem abre caminho para sistemas capazes de monitorar exercícios físicos em tempo real e fornecer *feedback* corretivo durante a prática.

Estudos exploram o uso de técnicas de visão computacional e estimativa de pose humana para análise automatizada de exercícios físicos. Essas abordagens utilizam bibliotecas de HPE para detectar pontos anatômicos do corpo humano e extrair métricas biomecânicas relacionadas ao movimento, como ângulos articulares, amplitude de movimento e padrões posturais. De acordo com Tharatipyakul, Srikaewsiew e Pongnumkul (2024), modelos baseados em *deep learning*, como OpenPose, MediaPipe e PoseNet, permitem capturar automaticamente *keypoints* do corpo humano a partir de imagens ou vídeos, possibilitando a análise da execução de exercícios físicos, a detecção de erros posturais e o fornecimento de *feedback* automático em aplicações de fitness, reabilitação e treinamento físico.

O sistema *Smart Physiotherapy*, desenvolvido por Hussain et al. (2024), utiliza a biblioteca MediaPipe para extração de marcos corporais e modelos de aprendizado de máquina baseados em *ensemble* para classificação de exercícios fisioterapêuticos. O estudo reportou acurácia de 98% na identificação de sete tipos de exercícios, com sensibilidade e especificidade superiores a 98%.

De forma semelhante, Sousa, Mauricio e Peres (2024) propuseram um sistema focado na análise de exercícios de calistenia utilizando estimativa de pose humana. O estudo destaca a viabilidade da modalidade como cenário para quantificar ângulos articulares e identificar desvios posturais em tempo real, sem a necessidade de sensores vestíveis. Os autores ressaltam que a integração entre visão computacional e calistenia viabiliza o desenvolvimento de ferramentas de monitoramento automatizado, promovendo maior autonomia e segurança na execução dos movimentos.

Outra abordagem foi apresentada por Ko, Nasridinov e Park (2024), que desenvolveram um sistema de correção postural em tempo real integrando as tecnologias YOLOv5 e MediaPipe. O sistema foi aplicado à análise do agachamento e obteve acurácia superior a 98% na identificação de

execuções corretas e incorretas do movimento.

Neste cenário, destaca-se o estudo de Rodriguez e Oliveira (2025), que investigou diferentes abordagens para análise de exercícios físicos utilizando visão computacional e aprendizado de máquina. O trabalho aplicou a biblioteca MediaPipe para estimativa de pose e utilizou redes *Long Short-Term Memory* (LSTM) para análise temporal dos movimentos, obtendo resultados promissores na classificação de exercícios.

Com todos os avanços observados, o trabalho de Rodriguez e Oliveira (2025) limitou-se a um conjunto reduzido de exercícios e não apresentou uma aplicação funcional em tempo real. De forma mais ampla, a literatura ainda apresenta limitações quando se considera o desenvolvimento de sistemas completos voltados ao acompanhamento da execução de exercícios físicos em cenários reais. Muitos estudos concentram-se apenas na classificação de movimentos ou na quantificação de parâmetros biomecânicos, enquanto outros são direcionados exclusivamente a contextos específicos, como fisioterapia ou treinamento com cargas externas.

O presente trabalho tem como objetivo desenvolver um sistema computacional para análise automatizada de exercícios físicos, com foco em calistenia e treinos domésticos. A proposta busca integrar a estimativa de pose humana à análise biomecânica para avaliar a qualidade da execução e fornecer *feedback* corretivo ao usuário em tempo real, contribuindo para uma prática autônoma mais segura e acessível.

Diante do crescimento das rotinas de exercícios sem supervisão e do consequente risco de lesões por execução inadequada, a visão computacional emerge como uma ferramenta crítica para democratizar o suporte técnico acessível. A carência de orientação qualificada em ambientes domésticos torna necessária a implementação de soluções que monitorem a biomecânica do movimento de forma automatizada, garantindo a segurança do praticante através de métricas como erro angular e estabilidade de *keypoints*.

Para fundamentar essa solução, o trabalho adota uma abordagem aplicada e descritiva, com análise prioritariamente qualitativa e teórica sobre a viabilidade técnica em tempo real. O estudo diferencia-se pela curadoria de uma base de dados híbrida para exercícios e pelo uso de uma arquitetura baseada em MediaPipe e redes LSTM, permitindo a validação conceitual da estabilidade temporal e da consistência dos ângulos articulares durante a prática dos exercícios.

O trabalho está organizado em quatro seções principais. A In-

trodução contextualiza a relevância da visão computacional na orientação de exercícios físicos, apresenta os fundamentos teóricos sobre calistenia e biomecânica do movimento, discute os trabalhos correlatos e define os objetivos da pesquisa. A seção de Materiais e Métodos descreve a construção da base de dados, a arquitetura dos modelos de classificação e o pipeline de inferência em tempo real. A seção de Resultados e Discussão apresenta os resultados obtidos no treinamento dos modelos, avalia o desempenho do sistema e discute as contribuições e limitações do trabalho. Por fim, a Conclusão sintetiza os principais achados e aponta direções para trabalhos futuros.

2 MATERIAIS E MÉTODOS

O desenvolvimento do sistema foi realizado com Python 3.11. As principais bibliotecas empregadas foram o MediaPipe 0.10.33, responsável pela estimativa de pose e extração de *keypoints* corporais; o TensorFlow 2.x com Keras, utilizado no treinamento e inferência dos modelos LSTM; o OpenCV 4.13, para captura e processamento de vídeo; o NumPy 2.4.4, para manipulação de *arrays* numéricos; o Pandas 3.0.2, para gerenciamento dos metadados da base de dados; e o Pillow, para renderização dos frames na interface gráfica. O ambiente de desenvolvimento foi executado em *hardware* composto por processador Intel Core i5-12450H, 16 GB de memória RAM DDR5 e placa de vídeo NVIDIA GeForce RTX 3050, com sistema operacional Microsoft Windows 11.

Foram selecionados exercícios de calistenia abrangendo movimentos dinâmicos com repetições e posições isométricas estáticas. A Tabela 1 apresenta os exercícios selecionados, sua classificação e o ângulo de câmera adotado.

Tabela 1 - Exercícios analisados, classificação e critérios de captura.

Exercício	Tipo	Âng. câmera	Principal erro detectável
Flexão (push-up)	Dinâmico	Lateral 90°	Quadril caindo, cotovelo aberto
Prancha (plank)	Estático	Lateral 90°	Quadril caindo ou empinando, tronco torto
Prancha lateral (side plank)	Estático	Frontal	Quadril caindo lateralmente
Posição de canoa (hollow body hold)	Estático	Lateral 90°	Joelho dobrado, pernas muito elevadas

Fonte: Elaborado pelo autor.

A Tabela 1 evidencia que a escolha do ângulo de câmera não foi arbitrária, mas definida conforme o erro postural mais relevante de cada exercício: a vista lateral favorece a leitura dos ângulos de joelho, quadril e alinhamento do tronco em flexão, prancha e posição de canoa, enquanto a vista frontal da prancha lateral privilegia a identificação de assimetrias entre os lados do corpo, conforme discutido por Baldinger, Reimer e Senner (2025).

Figura 1 - Quatro exercícios de calistenia analisados pelo sistema, com classificação e ângulo de câmera adotado para cada um.



Fonte: Elaborado pelo autor.

2.1 BASE DE DADOS

A base de dados foi construída a partir de vídeos de execuções corretas dos exercícios selecionados, coletados de fontes públicas e processados por um pipeline computacional desenvolvido para o projeto. Esta subseção descreve as fontes de coleta, o processo de verificação e triagem, a extração de características, as estratégias de augmentação e a geração sintética de amostras incorretas.

2.1.1 Coleta e triagem de vídeos

O ângulo de câmera foi padronizado por exercício com base na relevância biomecânica das informações capturadas: a vista lateral permite observar ângulos de joelho, quadril e alinhamento do tronco, enquanto a vista frontal é adequada para capturar simetria entre os lados e desloca-

mento lateral do quadril. Todos os vídeos coletados foram submetidos a um processo de triagem antes do processamento. Ao todo, foram coletados 328 vídeos, dos quais 23 foram descartados por apresentarem ângulo de câmera inadequado, oclusão severa de *keypoints* ou qualidade insuficiente de detecção, resultando em 305 vídeos em 30 fps aproveitados, distribuídos conforme a Tabela 2.

Tabela 2 - Distribuição dos vídeos coletados por exercício.

Exercício	Coletados	Descartados	Aproveitados	Duração média
Flexão (push-up)	93	23	70	≈ 5 s
Prancha (plank)	8	0	8	≈ 30 s
Prancha lateral (side plank)	13	0	13	≈ 30 s
Posição de canoa	37	0	37	≈ 10 s
Outro (UCF101)	214	0	214	≈ 3 s
Total	365	23	342	—

Fonte: Elaborado pelo autor.

A Tabela 2 mostra que a maior taxa de descarte concentrou-se na flexão de braço, exercício com maior volume de vídeos coletados, enquanto prancha, prancha lateral e a classe *outro* não apresentaram perdas, refletindo a qualidade da triagem aplicada antes do processamento.

A coleta foi híbrida, combinando *datasets* públicos e vídeos do YouTube. Para a flexão de braço, foram utilizados vídeos provenientes do *Kaggle Workout Fitness Video*, composto por vídeos reais de instrutores, totalizando 70 vídeos aproveitados após a triagem. Para prancha e prancha lateral não foram identificados *datasets* públicos adequados, sendo a coleta realizada exclusivamente via YouTube, com critério de seleção baseado no ângulo de câmera e na qualidade da detecção do MediaPipe, o que resultou em zero vídeos descartados para esses exercícios.

Para viabilizar a coleta dos vídeos do YouTube de forma eficiente, foi desenvolvida uma aplicação web auxiliar, executada localmente, que permite informar a URL do vídeo, selecionar o intervalo de tempo desejado para recorte e realizar o download diretamente na pasta de destino do projeto. A ferramenta foi implementada em *Python*, utilizando o módulo nativo `http.server` para disponibilizar uma interface *web* em *HTML* e *JavaScript* acessível pelo navegador, enquanto as operações de obtenção e processamento dos vídeos são delegadas, respectivamente, às bibliotecas `yt-dlp`, responsável pelo *download* a partir da URL informada com seleção de qualidade, e `FFmpeg`, empregada no recorte por intervalo de tempo sem recodificação (*stream copy*), o que preserva a qualidade original e re-

duz significativamente o tempo de processamento. A aplicação organiza os arquivos resultantes em diretórios nomeados conforme cada categoria de exercício, executa as tarefas de *download* e corte de forma assíncrona por meio de *threads* e fornece pré-visualização dos vídeos antes do recorte, eliminando a necessidade de processamento manual de edição e organização dos arquivos, e agilizando significativamente o processo de curadoria do *dataset*.

A base de dados é composta por três classes distintas, cada uma com finalidade específica no treinamento dos modelos. A classe *correto* reúne sequências de execuções biomecânicas adequadas de cada exercício, servindo como referência positiva para o aprendizado. A classe *incorreto* contém sequências geradas sinteticamente a partir de perturbações controladas dos ângulos articulares das amostras corretas, representando desvios posturais comuns em cada exercício. A classe *outro* corresponde a sequências de movimentos que não pertencem a nenhum dos exercícios analisados, como pessoas em pé, sentadas, caminhando ou realizando atividades cotidianas.

A classe *outro* foi criada especificamente para treinar o detector global, cujo papel é identificar quando o usuário não está executando o exercício selecionado, problema conhecido na literatura como classificação fora da distribuição de treino (*out-of-distribution*). Sem essa classe, os modelos avaliadores tendem a emitir classificações biomecânicas com alta confiança mesmo para poses completamente alheias ao exercício, tornando o *feedback* inválido. Para a classe *outro*, foram utilizados 214 vídeos provenientes do *dataset* público UCF101, composto por registros de pessoas em posições verticais, como caminhada e corrida, todos aproveitados sem necessidade de descarte.

A estratégia de coleta considerou a natureza de cada exercício: para exercícios dinâmicos, a diversidade é obtida por meio de vídeos de pessoas diferentes, com variações de ritmo e amplitude; para exercícios estáticos, a diversidade é complementada pela duração do vídeo, uma vez que microoscilações e ajustes posturais ao longo do tempo geram variações naturais entre as janelas temporais extraídas.

2.1.2 Ferramentas de verificação e diagnóstico

Todos os vídeos foram verificados individualmente antes do processamento por meio de um conjunto de ferramentas de diagnóstico desenvolvidas especificamente para o projeto. Essas ferramentas cobrem três

etapas distintas do pipeline: inspeção visual dos vídeos, verificação numérica dos dados gerados e análise da qualidade do particionamento entre treino e teste.

Na etapa de inspeção visual, a ferramenta sobrepõe em tempo real o esqueleto detectado pelo MediaPipe sobre os frames do vídeo, permitindo confirmar o ângulo de câmera, a cobertura dos 33 pontos corporais e a estabilidade da detecção ao longo do tempo. Vídeos com detecção instável, ângulo inadequado ou oclusão severa foram descartados nessa etapa.

Na etapa de verificação numérica, após o processamento dos vídeos, as ferramentas permitiram inspecionar os arrays gerados, confirmando que os valores de coordenadas normalizadas e ângulos articulares eram consistentes com os valores biomecânicos esperados para cada exercício. Além disso, foi possível comparar a distribuição dos ângulos articulares entre amostras corretas e incorretas, confirmando que a perturbação sintética gerava separação estatística suficiente entre as classes antes do treinamento.

Na etapa de análise do particionamento, as ferramentas verificaram a distribuição dos vídeos entre os conjuntos de treino e teste, garantindo que vídeos com grande volume de amostras não dominassem o conjunto de avaliação. Essa análise foi especialmente relevante para a prancha lateral, que apresentava um vídeo com 620 amostras, cerca de 78% do conjunto de teste padrão, problema identificado e corrigido com a escolha de um particionamento mais equilibrado entre os vídeos disponíveis.

2.1.3 Extração de características

A extração de características foi realizada com o modelo completo do MediaPipe Pose Landmarker, operando no modo `VIDEO` com *timestamps* monotônicos. O MediaPipe Pose é uma biblioteca desenvolvida pelo Google que aplica redes neurais profundas para detectar 33 pontos anatômicos do corpo humano a partir de vídeos RGB comuns, sem necessidade de sensores especializados ou marcadores físicos Lan et al. (2023). Os 33 pontos cobrem as principais articulações do corpo, incluindo ombros, cotovelos, pulsos, quadris, joelhos e tornozelos, e cada ponto é representado por coordenadas x e y normalizadas no intervalo $[0, 1]$ em relação às dimensões do *frame*, além de um atributo de visibilidade indicando a confiança da detecção. *Frames* em que a média de visibilidade dos 33 *keypoints* ficou abaixo de 0,5 foram descartados automaticamente, evitando a incor-

poração de detecções de baixa qualidade à base.

Além das coordenadas dos keypoints, foram calculados nove ângulos articulares por frame, obtidos por meio do produto escalar entre os vetores formados por segmentos corporais consecutivos. Essa abordagem, adotada por Mercadal-Baudart et al. (2024) e Ko, Nasridinov e Park (2024), torna as métricas parcialmente invariantes a variações de distância da câmera e orientação do corpo. A Tabela 3 apresenta os ângulos calculados e os pontos utilizados.

Tabela 3 - Ângulos articulares calculados por frame.

Ângulo	Pontos (A – B – C)	Articulação
cotovelo_esq	kp15 – kp13 – kp11	Cotovelo esquerdo
cotovelo_dir	kp16 – kp14 – kp12	Cotovelo direito
joelho_esq	kp23 – kp25 – kp27	Joelho esquerdo
joelho_dir	kp24 – kp26 – kp28	Joelho direito
quadril_esq	kp11 – kp23 – kp25	Quadril esquerdo
quadril_dir	kp12 – kp24 – kp26	Quadril direito
alinhamento	kp11 – kp23 – kp27	Alinhamento ombro-quadril-tornozelo
ombro_esq	kp13 – kp11 – kp23	Ombro esquerdo
ombro_dir	kp14 – kp12 – kp24	Ombro direito

Fonte: Elaborado pelo autor.

Os nove ângulos apresentados na Tabela 3 cobrem as principais articulações envolvidas nos exercícios selecionados, como cotovelos, joelhos, quadris e o alinhamento ombro-quadril-tornozelo, fornecendo a base biomecânica tanto para o treinamento dos modelos quanto para as regras de detecção de erro descritas na Seção 2.2, em linha com a abordagem adotada por Mercadal-Baudart et al. (2024) e Ko, Nasridinov e Park (2024).

2.1.4 Formato e estrutura do dataset

Cada amostra da base é armazenada em um arquivo NumPy (.npy) com estrutura de formato (30, 108), onde 30 representa o número de frames por sequência e 108 o total de atributos por frame, 99 valores de coordenadas dos keypoints concatenados com os 9 ângulos articulares. As sequências foram extraídas por meio de uma janela deslizante de 30 frames com passo de 15 frames, resultando em sobreposição de 50% entre janelas consecutivas. A 30 frames por segundo, cada sequência corresponde a aproximadamente um segundo de movimento. Os arquivos foram indexados por um arquivo de metadados no formato CSV, registrando para cada amostra o caminho do arquivo, o exercício, a classe, o identificador do vídeo de origem e a fonte de coleta.

2.1.5 Augmentação e geração sintética de incorretos

Para ampliar o volume da base sem necessidade de coleta adicional, foi aplicado espelhamento horizontal sobre todas as sequências geradas, invertendo o eixo x dos keypoints e dobrando o número de amostras por vídeo. Os ângulos articulares não são afetados por essa operação, pois são invariantes à orientação do corpo.

As amostras da classe incorreto foram geradas sinteticamente a partir da perturbação controlada dos ângulos articulares de sequências corretas. Para cada exercício, os ângulos mais relevantes foram deslocados com intensidade variável: entre 10° e 20° para flexão e prancha lateral, entre 25° e 35° para prancha e entre 30° e 40° para a posição de canoa, valores maiores adotados para representar erros visualmente evidentes, dado que para esse exercício não foram identificados estudos científicos com limites angulares validados, sendo os critérios definidos empiricamente a partir de vídeos instrucionais que demonstram explicitamente execuções incorretas, valor maior adotado para compensar o efeito de perspectiva da câmera, que faz com que o ângulo de alinhamento de execuções corretas de prancha fique naturalmente próximo de 160° em vez de 180° . O balanço entre as classes correto e incorreto foi mantido em proporção 1:1 por exercício. Essa abordagem está alinhada à metodologia de Rodriguez e Oliveira (2025) e é reconhecida na literatura como *augmentação sintética de poses*.

2.2 ARQUITETURA DO SISTEMA

A Figura 2 sintetiza o fluxo completo do sistema, desde a coleta dos vídeos até a entrega do *feedback* corretivo ao usuário, ilustrando a integração entre as etapas descritas nas subseções seguintes.

O sistema adota uma arquitetura de inferência em cascata composta por dois estágios independentes. O primeiro estágio consiste em um modelo detector global, treinado para distinguir se o conteúdo capturado pela câmera corresponde a um exercício de calistenia ou a qualquer outra posição, denominada classe outro. O segundo estágio consiste nos modelos avaliadores, um por exercício, responsáveis por classificar a execução como correta ou incorreta. Essa separação em dois estágios resolve o problema de classificações fora da distribuição de treino: quando o usuário não está na posição do exercício selecionado, o detector rejeita a entrada e nenhuma avaliação biomecânica é realizada, evitando *feedbacks* incorretos.

Figura 2 - *Pipeline* de processamento em seis etapas: coleta, triagem e validação, extração de *keypoints*, geração de amostras, treino em cascata e inferência com *feedback* em tempo real.



Fonte: Elaborado pelo autor.

O particionamento entre treino e teste foi realizado por vídeo de origem, por meio da técnica *GroupShuffleSplit*, na proporção de 80% para treino e 20% para teste. Dessa proporção de treino, outros 20% foram reservados para validação interna durante o treinamento, resultando em aproximadamente 64% dos dados destinados ao treino efetivo, 16% à validação e 20% ao teste final. Essa estratégia garante que todas as sequências de um mesmo vídeo pertençam exclusivamente a um único conjunto, evitando o vazamento de dados entre amostras originais e espelhadas provenientes do mesmo vídeo de origem.

A LSTM foi selecionada pela sua capacidade de capturar dependências temporais em sequências de dados, relacionando *frames* consecutivos de uma execução, o que não é possível com modelos que analisam *frames* de forma isolada, conforme destacado por Nguyen et al. (2024). Diferentemente da abordagem comum na literatura, a LSTM foi adotada também para os exercícios isométricos, com base na premissa de que micro-oscilações e ajustes posturais contínuos geram padrões temporais distintos entre execuções estáveis e instáveis.

A aplicação de *feedback* em tempo real consome os modelos treinados sem acesso ao código de treinamento. A interface foi construída com Tkinter e OpenCV, permitindo o uso tanto de *webcam* quanto de arquivos de vídeo como fonte de entrada, o que viabiliza testes com vídeos de terceiros sem necessidade de nova gravação. A captura por *webcam* opera

a 30 *frames* por segundo com resolução de 1280×720 pixels. O *pipeline* de inferência em tempo real é idêntico ao utilizado no treinamento: para cada *frame* válido, os 33 *keypoints* detectados pelo MediaPipe são concatenados com os 9 ângulos articulares, formando um vetor de 108 *features*. Esses vetores são acumulados em um *buffer* circular de 30 *frames* e, a cada 15 *frames* novos, o *array* (1, 30, 108) é apresentado aos modelos.

Para garantir estabilidade visual do *feedback* durante o exercício, foram implementadas três camadas de suavização. A primeira aplica um filtro de média móvel exponencial (EMA) com fator $\alpha = 0,25$ sobre os ângulos articulares utilizados na exibição e na detecção de erros posturais, sem afetar os valores brutos apresentados aos modelos. A segunda aplica EMA com $\alpha = 0,30$ sobre os *landmarks* para reduzir o tremor visual do esqueleto. A terceira implementa um mecanismo de histerese com votação no avaliador: para alternar entre correto e incorreto após a primeira inferência, o modelo precisa emitir votos consecutivos na nova classe por pelo menos duas inferências seguidas, equivalente a aproximadamente um segundo de sinal contínuo. Uma zona cinza definida entre 0,40 e 0,60 na saída sigmoid mantém a classificação anterior sem atualizar a confiança exibida.

Quando o avaliador classifica a execução como incorreta, o sistema identifica o erro postural específico por meio de regras biomecânicas aplicadas sobre os ângulos suavizados. Cada regra calcula o desvio em graus em relação a um limiar estabelecido com base na literatura biomecânica, e a regra com maior desvio é reportada ao usuário. As Tabelas 4 e 5 apresentam as regras implementadas para cada exercício.

Tabela 4 - Regras biomecânicas de detecção de erro postural: Flexão e prancha

Exercício	Erro	Condição	Instrução de correção
Flexão	Quadril caindo	alinhamento < 160	Eleve o quadril para alinhar com os ombros
	Quadril empinando	alinhamento > 190	Abaixe o quadril para alinhar com os ombros
	Cotovelo aberto	cotovelo_esq > 160 e cotovelo_dir > 160	Mantenha os cotovelos próximos ao corpo
Prancha	Quadril caindo	alinhamento < 160	Contraia o core e eleve o quadril
	Quadril empinando	alinhamento > 190	Abaixe o quadril até alinhar com o tronco
	Tronco rodado	quadril_esq - quadril_dir > 15	Mantenha os quadris paralelos ao chão
	Cabeça baixa	min(ombro_esq, ombro_dir) < 50	Mantenha a cabeça alinhada com a coluna

Fonte: Elaborado pelo autor.

Tabela 5 - Regras biomecânicas de detecção de erro postural: Prancha lateral e posição de canoa

Exercício	Erro	Condição	Instrução de correção
Prancha lateral	Quadril caindo	$\text{quadril_esq} < 150$ e $\text{quadril_dir} < 150$	Eleve o quadril lateralmente
	Corpo desalinhado	$ \text{ombro_esq} - \text{ombro_dir} > 20$	Alinhe os ombros verticalmente
Posição de canoa	Joelho dobrado	$(\text{joelho_esq} + \text{joelho_dir}) / 2 < 140$	Estenda completamente as pernas
	Pernas elevadas	$(\text{quadril_esq} + \text{quadril_dir}) / 2 > 160$	Abaixe as pernas mantendo a lombar no chão
	Cabeça apoiada	$\min(\text{ombro_esq}, \text{ombro_dir}) < 50$	Eleve a cabeça com o queixo próximo ao peito
	Braços dobrados	$\min(\text{cotovelo_esq}, \text{cotovelo_dir}) < 150$	Estenda os braços acima da cabeça

Fonte: Elaborado pelo autor.

As regras apresentadas nas Tabelas 4 e 5 traduzem os limiares angulares discutidos na fundamentação teórica em condições objetivas de verificação, permitindo que o sistema não apenas classifique a execução como incorreta, mas também identifique o desvio postural específico, como quadril caindo ou cotovelo aberto, priorizando sempre a regra de maior magnitude de erro detectado.

2.3 APRESENTAÇÃO DO FEEDBACK

O *feedback* é apresentado diretamente sobre o vídeo por meio de um *banner* no topo do *frame*, codificado por cor: azul durante o preenchimento inicial do *buffer*, cinza quando o detector classifica a entrada como fora de posição, verde para execução correta e vermelho para execução incorreta. Neste último caso, o nome do erro e a magnitude do desvio em graus são exibidos junto à instrução de correção. Essa escolha de apresentação sobreposta ao vídeo foi adotada com base na premissa de que, durante o exercício, o usuário acompanha a própria imagem na tela e não um painel lateral com valores numéricos.

3 RESULTADOS E DISCUSSÃO

Os resultados obtidos são apresentados em dois níveis: o desempenho dos modelos avaliado sobre o conjunto de teste durante o treinamento, e os resultados da validação funcional da aplicação fornecendo *feedback* em tempo real.

As métricas utilizadas foram acurácia, *precision*, *recall* e *F1-score*, calculadas por classe. A escolha do *F1-score* como métrica complementar à acurácia se justifica pelo contexto da aplicação: falsos positi-

vos geram alertas de erro desnecessários, comprometendo a experiência do usuário, enquanto falsos negativos deixam passar erros posturais reais sem *feedback*, o que é mais crítico do ponto de vista da segurança do praticante. A Tabela 6 apresenta os resultados consolidados.

Tabela 6 - Resultados do treino dos modelos avaliados no conjunto de teste.

Modelo	Acurácia	Classe	Precision	Recall	F1
Detector	94,0%	exercício	0,99	0,92	0,96
		outro	0,83	0,98	0,90
Flexão	91,5%	correto	0,92	0,91	0,91
		incorreto	0,91	0,92	0,91
Prancha	100%	correto	1,00	1,00	1,00
		incorreto	1,00	1,00	1,00
Prancha lateral	83,1%	correto	0,81	0,86	0,84
		incorreto	0,85	0,80	0,83
Posição de canoa	99,8%	correto	1,00	1,00	1,00
		incorreto	1,00	1,00	1,00

Fonte: Elaborado pelo autor.

O detector global atingiu 94,0% de acurácia, com *recall* de 0,98 para a classe outro, indicando que o sistema identifica com consistência quando o usuário não está na posição do exercício selecionado. Esse resultado é especialmente relevante considerando que a posição de canoa, por ser realizada deitado de costas, apresenta padrão de *keypoints* distinto dos demais exercícios e mais próximo de algumas poses da classe outro, o que exigiu o aumento do volume de amostras dessa classe durante o desenvolvimento para manter a qualidade da detecção.

Entre os avaliadores, a prancha e a posição de canoa atingiram acurácia de 100% e 99,8%, respectivamente. Esses resultados elevados refletem a consistência temporal característica dos exercícios isométricos, cujos padrões de *keypoints* variam pouco entre repetições e indivíduos, combinada com a separação angular acentuada entre as classes correto e incorreto gerada pela perturbação sintética. É importante destacar, contudo, que o número reduzido de vídeos de origem para esses exercícios, especialmente a prancha, com apenas 8 vídeos, limita a representatividade do conjunto de teste, o que pode contribuir para resultados artificialmente elevados que não necessariamente se generalizam para toda a variabilidade possível de execução em condições reais. Essa hipótese é parcialmente corroborada pelos resultados da validação funcional, onde a prancha atingiu 92,3%, resultado sólido, porém inferior aos 100% do conjunto de teste, sugerindo que a avaliação com vídeos externos captura melhor a

variabilidade real do exercício.

A validação funcional da aplicação foi realizada por meio de uma avaliação em lote com vídeos externos ao *dataset* de treinamento, totalizando 77 vídeos, dos quais 3 foram descartados por insuficiência na detecção de pose pelo MediaPipe, todos pertencentes à categoria outro, onde condições de iluminação desfavoráveis comprometeram a visibilidade dos *keypoints*, resultando em 74 vídeos efetivamente processados e 1.228 janelas inferidas pelo sistema completo. Para cada exercício foram utilizados vídeos de execuções corretas e incorretas de pessoas não presentes no *dataset* de treinamento, e o veredito de cada vídeo foi determinado pelo voto majoritário entre suas janelas, considerando o *pipeline* completo de detector seguido de avaliador. A Tabela 7 apresenta os resultados obtidos.

Tabela 7 - Resultados da validação funcional da aplicação em tempo real.

Exercício	Acerto em corretos	Acerto em incorretos	Acurácia
Flexão	8/8 (100,0%)	9/11 (81,8%)	89,5%
Prancha	7/7 (100,0%)	5/6 (83,3%)	92,3%
Prancha lateral	7/9 (77,8%)	6/8 (75,0%)	76,5%
Posição de canoa	8/9 (88,9%)	9/10 (90,0%)	89,5%
Outro	—	6/6 (100,0%)	100,0%
Média			87,8%

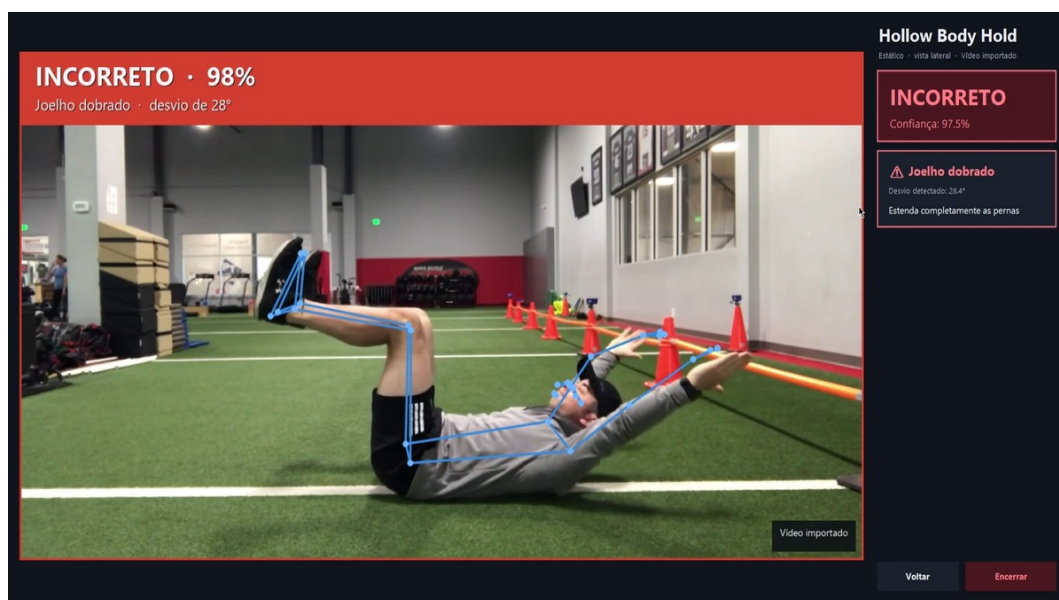
Fonte: Elaborado pelo autor.

A acurácia geral da aplicação em tempo real atingiu 87,8%, considerando o *pipeline* completo de detector e avaliador sobre vídeos externos ao *dataset* de treinamento. O detector global classificou corretamente todos os 6 vídeos da categoria outro como poses fora do escopo, demonstrando eficácia na rejeição de entradas fora da distribuição de treino. Entre os exercícios, a prancha obteve o melhor resultado com 92,3%, seguida da flexão e da posição de canoa com 89,5% cada. A prancha lateral apresentou o menor resultado na validação funcional, 76,5%, valor consistente com o desempenho observado durante o treinamento e que reforça a necessidade de ampliar o volume de vídeos de origem para esse exercício em trabalhos futuros.

Os resultados dos exercícios incorretos foram em geral inferiores aos corretos, o que é esperado considerando que as amostras incorretas do *dataset* de treinamento foram geradas sinteticamente, enquanto os vídeos de validação continham execuções incorretas reais com padrões de erro naturalmente mais variados. A exceção foi a posição de canoa, cujos incorretos atingiram 90,0%, superior ao acerto nos corretos, indicando

que os erros sintéticos gerados para esse exercício representaram bem os padrões de erro observados na prática.

Figura 3 - Exemplo de *feedback* em tempo real da aplicação para a posição de canoa (*hollow body hold*): execução classificada como incorreta com 97,5% de confiança, erro identificado como joelho dobrado com desvio de 28,4°.



Fonte: Elaborado pelo autor.

A Figura 3 ilustra o comportamento da aplicação durante a validação funcional: o *banner* vermelho no topo do *frame* exibe a classificação, a confiança do modelo, o nome do erro detectado, o desvio angular calculado e a instrução de correção correspondente, todos sobrepostos diretamente ao vídeo do usuário.

A inclusão de exercícios isométricos estáticos representa um diferencial em relação ao trabalho de Ko, Nasridinov e Park (2024), que desenvolveram um sistema de correção postural em tempo real para o agachamento, obtendo acurácia superior a 98%, porém restrito a um único exercício dinâmico e ao contexto de *powerlifting* com cargas externas.

Em relação ao trabalho de Rodriguez e Oliveira (2025), que adota a mesma combinação MediaPipe e LSTM, o particionamento por divisão aleatória simples adotado naquele estudo não controla quais vídeos compõem cada conjunto. Como cada vídeo gerou amostras originais e espelhadas, versões do mesmo vídeo podem ter ficado simultaneamente no treino e no teste, o que pode ter inflado a acurácia reportada. No presente trabalho, o uso do *GroupShuffleSplit* por vídeo de origem garante uma avaliação

mais conservadora e metodologicamente mais confiável da capacidade de generalização dos modelos para novos usuários. Além disso, aquele estudo limitou-se a três exercícios e não apresentou uma aplicação funcional em tempo real, aspectos que o presente trabalho buscou avançar.

Uma limitação identificada refere-se à fundamentação dos limiares angulares utilizados nas regras de detecção de erros posturais. Para a maioria dos exercícios analisados, não foram identificados estudos científicos que estabeleçam valores angulares validados experimentalmente para a classificação de execuções corretas e incorretas em contexto de calistenia, conforme discutido na Seção Materiais e Métodos.

4 CONCLUSÃO

O presente trabalho teve como objetivo desenvolver um sistema computacional para análise biomecânica de exercícios de calistenia em tempo real, integrando estimativa de pose humana, classificação por redes LSTM e *feedback* corretivo ao usuário. O objetivo foi atingido com a entrega de uma aplicação funcional capaz de classificar a qualidade da execução e identificar o erro postural específico com instrução de correção em tempo real, sem necessidade de sensores especializados ou supervisão presencial.

Entre as contribuições do trabalho, destacam-se a inclusão de exercícios isométricos estáticos, modalidade pouco explorada na literatura de análise automatizada de calistenia e não contemplada por trabalhos correlatos voltados exclusivamente a exercícios dinâmicos com cargas externas, e o desenvolvimento de um *pipeline* completo que integra coleta de dados, treinamento e inferência em uma ferramenta acessível para a prática autônoma de exercícios com suporte a múltiplos exercícios simultaneamente. Uma lição relevante identificada durante o desenvolvimento foi que a diversidade dos vídeos de origem exerce maior influência sobre o desempenho dos modelos do que o volume absoluto de amostras, observação que deve orientar futuras expansões do sistema.

Como trabalhos futuros, recomenda-se a validação dos limiares angulares adotados com especialistas em biomecânica e educação física, a ampliação do repertório de exercícios analisados, e a investigação de abordagens para cenários com múltiplas pessoas no campo de visão, limitação identificada no sistema atual.

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL

Os autores declaram que ferramentas de Inteligência Artificial generativa foram utilizadas exclusivamente como apoio auxiliar à elaboração deste trabalho, incluindo assistência na revisão linguística, estruturação textual e organização de ideias. Nenhuma ferramenta de IA foi utilizada para substituição da autoria intelectual, análise científica, interpretação dos resultados ou tomada de decisões acadêmicas. A responsabilidade integral pelo conteúdo, originalidade, veracidade das informações e conformidade com os princípios de integridade científica permanece integralmente atribuída aos autores.

REFERÊNCIAS

- BALDINGER, M.; REIMER, L. M.; SENNER, V. **Influence of the camera viewing angle on openpose validity in motion analysis**. *Sensors*, v. 25, n. 3, p. 799. Acesso em: 30 set. 2025. 2025, Disponível em: <<https://www.mdpi.com/1424-8220/25/3/799>>.
- BONILLA, D. A. et al. **Exercise Selection and Common Injuries in Fitness Centers: A Systematic Integrative Review and Practical Recommendations**. 2022. 12710 p. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/36232010/>>. Acesso em: 26 ago. 2025.
- CAO, Z. et al. **Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields**. 2017. 7291–7299 p. Disponível em: <https://openaccess.thecvf.com/content_cvpr_2017/papers/Cao_Realtime_Multi-Person_2D_CVPR_2017_paper.pdf>. Acesso em: 25 ago. 2025.
- HUSSAIN, A. et al. **Smart physiotherapy: Advancing arm-based exercise classification with posenet and ensemble models**. *Sensors*, v. 24, n. 4, p. 1056. Acesso em: 26 out. 2025. 2024, Disponível em: <<https://www.mdpi.com/1424-8220/24/19/6325>>.
- IVERSEN, V. M.; BØ, K.; FIMLAND, M. S. **Calisthenics training and its effects on physical fitness: A systematic review**. *Sports Medicine - Open*, v. 9, n. 1, p. 56. Acesso em: 27 out. 2025. 2023, Disponível em: <<https://sportsmedicine-open.springeropen.com/articles/10.1186/s40798-023-00543-1>>.
- KO, Y.-M.; NASRIDINOV, A.; PARK, S.-H. **Real-time ai posture correction for powerlifting exercises using yolov5 and mediapipe**. *IEEE Access*, v. 12, p. 195830–195853. Acesso em: 25 nov. 2025. 2024, Disponível em: <<https://ieeexplore.ieee.org/document/10798440>>.
- LAN, G. et al. **Vision-based human pose estimation via deep learning: A survey**. *arXiv*, abs/2308.13872. Acesso em: 05 nov. 2025. 2023, Disponível em: <<https://arxiv.org/abs/2308.13872>>.

MAUSE, R. et al. **Biomechanical analysis of bodyweight squats: Influence of stance width and depth on joint loading and muscle activation.** *Frontiers in Sports and Active Living*, v. 6, p. 1423285. Acesso em: 27 out. 2025. 2024, Disponível em: <<https://pmc.ncbi.nlm.nih.gov/articles/PMC10987311/>>.

MERCADAL-BAUDART, C. et al. **Exercise quantification from single camera view markerless pose estimation: joint angles for sports physiotherapy.** *Heliyon*, v. 10. Acesso em: 30 set. 2025. 2024, Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2405844024036272>>.

NGUYEN, S. M. et al. **A medical low-back pain physical rehabilitation dataset for human body movement analysis.** *ArXiv preprint, CHRU Brest*. ArXiv:2407.00521. Acesso em: 2 mar. 2025. 2024, Disponível em: <<https://arxiv.org/pdf/2407.00521>>.

RODRIGUEZ, P. B.; OLIVEIRA, M. **Comparação de técnicas de aprendizado de máquina para análise de movimentos biomecânicos em exercícios físicos.** Universidade do Extremo Sul Catarinense, 2025. Disponível em: <<http://repositorio.unesc.net/handle/1/11893>>. Acesso em: 11 set. 2025.

SOUSA, M. O. de; MAURICIO, C. R. M.; PERES, F. F. F. **Análise e orientação de postura nos exercícios de calistenia usando estimativa de pose humana.** In: SBC. *Anais do XXI Congresso Latino-Americano de Software Livre e Tecnologias Abertas (Latinoware 2024)*. 2024. 2024. p. 368–371. Acesso em: 19 mar. 2025. Disponível em: <<https://sol.sbc.org.br/index.php/latinoware/article/view/31553>>.

SZAJKOWSKI, S. et al. **Risk Factors for Injury in CrossFit®—A Retrospective Analysis.** 2023. 2211 p. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/36767578/>>. Acesso em: 26 ago. 2025.

THARATIPYAKUL, A.; SRIKAEWSIEW, T.; PONGNUMKUL, S. **Deep learning-based human body pose estimation in providing feedback for physical movement: A review.** *Heliyon*, v. 10, n. 6, p. e27452. Acesso em: 16 mar. 2026. 2024, Disponível em: <[https://www.cell.com/heliyon/fulltext/S2405-8440\(24\)12620-5](https://www.cell.com/heliyon/fulltext/S2405-8440(24)12620-5)>.

WEI, S. et al. **Lower-limb muscle adaptation to progressive bodyweight squat training: a controlled experimental study.** *Scientific Reports*, v. 13, p. 40319. Acesso em: 5 nov. 2025. 2023, Disponível em: <<https://www.nature.com/articles/s41598-023-40319-x>>.

World Health Organization. **WHO Guidelines on Physical Activity and Sedentary Behaviour.** Geneva: World Health Organization, 2020. Disponível em: <<https://www.who.int/publications/i/item/9789240015128>>. Acesso em: 25 ago. 2025.