

# UMA METODOLOGIA DE VISÃO COMPUTACIONAL PARA ANÁLISE DE MOVIMENTOS FISIOTERÁPICOS

Gustavo Garcia Bagio<sup>1</sup>, Paulo Sergio Silva Rodrigues<sup>2</sup>

<sup>1,2</sup> *Ciência da Computação, Centro Universitário FEI*

*unifgbagio@fei.edu.br psergio@fei.edu.br*

**Resumo:** A reabilitação física após tratamento de patologias é desafiadora para médicos, pacientes e familiares. Ainda que existam custos elevados em equipamentos e consultoria, avanços em Visão Computacional e Aprendizado de Máquina sugerem soluções mais econômicas. O monitoramento de pacientes com inteligência artificial, especialmente usando sinais visuais, é pouco explorado. Este estudo desenvolve um sistema com câmera RGB para rastrear articulações em tempo real, melhorando a fisioterapia.

## 1. Introdução

Recentemente, testemunhamos avanços significativos em inteligência artificial, incluindo o uso de aprendizado de máquina para tarefas como regressão, classificação [1] e redes neurais profundas para reconhecimento de fala [2]. A popularização dessa área deve-se à sua aplicação em diversos campos, desde engenharia até entretenimento. Contudo, um fator essencial tem sido o uso de tecnologias de baixo custo, como câmeras e sensores, em modelos preditivos que empregam sinais digitais como entrada.

A área da saúde, particularmente, tem se beneficiado consideravelmente da inteligência artificial. As aplicações incluem tratamentos de câncer com Redes Neurais Convolucionais [3], prevenção de quedas [4], tratamento de doenças crônicas [5] e reabilitação pós-AVC [6]. Atualmente, muitos tratamentos adotam abordagens híbridas, combinando o expertise de profissionais da saúde com algoritmos de aprendizado de máquina [6]. Especificamente para reabilitação pós-AVC, a monitorização dos exercícios, facilitada por sensores e câmeras, tem permitido que profissionais avaliem a execução dos exercícios remotamente, oferecendo mais eficiência e ampliando o acesso a tratamentos de alta qualidade em áreas remotas [7].

Porém, existem desafios a serem enfrentados, como a falta de distinção da IA entre AVCs graves e leves e a necessidade de avaliações mais detalhadas durante atividades reabilitativas. Adicionalmente, trabalhos recentes [8,9] apresentaram propostas promissoras utilizando Visão Computacional para avaliar pacientes durante exercícios, mas ainda dependem de hardware pouco acessível e enfrentam dificuldades relacionadas à proteção de dados e à formação de bases de dados adequadas.

Diante desses desafios, este projeto visa introduzir uma nova metodologia baseada em câmeras RGBD simples, visando sua aplicação em tempo real e com baixo custo. Paralelamente, propomos uma estratégia para geração de bases de dados de treinamento. O objetivo é automatizar o monitoramento de exercícios reabilitativos, identificando a performance e o tipo de

exercício realizado, contribuindo assim para a telemedicina, especialmente em regiões como o Brasil, com vastas dimensões e desigualdades sociais. Esta Iniciação Científica dá sequência ao trabalho realizado na tese de mestrado de João Aurélio Francisco Junior [10], ex-aluno do Centro Universitário FEI.

## 2. Metodologia

A metodologia proposta neste projeto é composta de nove etapas principais. O objetivo da metodologia é utilizar câmeras RGB que capturam imagens dos exercícios físicos realizados por pessoas, que são traduzidas para valores cinemáticos e então armazenados em bancos de ações. Tais bancos são utilizados para treinamento dos módulos de "Detecção" e "Medida". Como saída desses módulos, há uma resposta visual em forma de texto indicando para a pessoa as articulações que demandam correções durante o exercício.

A metodologia inicia na "Etapa (A): Exercícios", quando a pessoa realiza um conjunto de exercícios para membros inferiores e uma câmera RGB na "Etapa (B): Câmera RGB" captura seus movimentos. Na "Etapa (C): OpenPose", uma biblioteca em Python chamada OpenPose que utiliza o *TensorFlow*, é utilizada para identificar em cada *frame* as articulações do corpo. Na "Etapa (D): Extração de características cinemáticas", são capturados dados cinemáticos das articulações construídas na etapa anterior. Os dados cinemáticos aqui utilizados são os ângulos entre as articulações da extremidade superior e inferior do corpo da pessoa que está executando os exercícios. Da "Etapa (D)" para as "Etapa (E): Banco de Ações" e "Etapa (F): Banco de Ações + Oscilações", ocorre o armazenamento dos movimentos da pessoa no Banco de Ações utilizado para o treinamento dos módulos de "Detecção" e "Medida". Para treinar a rede, foi utilizado um Banco de Ações que armazena os dados cinemáticos capturados *frame-to-frame*. O Banco de Ações foi construído de forma manual. Na "Etapa (F)", com o objetivo de ampliar as ações da etapa anterior, foram adicionadas novas oscilações de movimentos. Essa ampliação permite que o módulo de "Medida" consiga detectar os exercícios feitos de forma incorreta.

Logo após a construção dos bancos de ações, os módulos de "Detecção" e "Medida" da rede de módulos neurais foram treinados. O módulo da "Etapa (G): Detecção" foi treinado para reconhecer a ação executada pela pessoa e identificá-la no Banco de Ações. Após essa identificação, a "Etapa (H): Medida" tem como entrada os vetores de feature da ação da pessoa junto à ação correta no Banco de Ações. Como saída, na "Etapa (I): Resposta Visual da Ação e Correção", foram

apresentados: um vídeo sobrepondo a imagem da pessoa; o tipo de exercício que está sendo realizado; e quais articulações estão em posições erradas no decorrer da atividade.

### 3. Resultados Parciais

A primeira etapa que foi concluída refere-se à Coleta de Dados. A fim de obter os dados necessários para o desenvolvimento do projeto, foram utilizadas três câmeras web, especificamente o modelo *Logitech C922 Pro Stream*. Essas câmeras foram escolhidas devido à sua capacidade de capturar vídeos com uma alta taxa de quadros, chegando a sessenta quadros por segundo.

A escolha por três câmeras se justifica pela necessidade de registrar os vídeos de voluntários realizando exercícios fisioterápicos em três perspectivas distintas: Lado Esquerdo, Frontal e Lado Direito. Inicialmente, contamos com a participação de 10 voluntários, todos alunos da instituição, que realizaram os exercícios propostos no projeto original e também um exercício adicional, o Agachamento unipodal com ambas as pernas.

A coleta desse conjunto inicial de dados tem o propósito de nos permitir conduzir uma série de testes. Estes testes são essenciais tanto para determinar a quantidade de dados que será necessária para o treinamento adequado da rede neural, quanto para avaliar a acurácia do algoritmo de Pose Estimation, especificamente o *OpenPose*.

Conforme o desenvolvimento do projeto progrediu, identificou-se a necessidade de uma etapa adicional: avaliar a precisão do algoritmo *OpenPose* na extração de dados que seriam usados para avaliar os exercícios fisioterápicos, visto que pequenas variações na posição detectada das articulações podem resultar em alterações significativas na avaliação dos exercícios.

Para avaliar a acurácia do *OpenPose*, foram capturados aproximadamente 10.000 frames de voluntários em posição estática para cada uma das três perspectivas (Lado Esquerdo, Frontal e Lado Direito), totalizando cerca de 30.000 frames. Estes frames foram processados pelo algoritmo *OpenPose* para estimar a pose. Em seguida, desenvolveu-se um algoritmo auxiliar destinado a tratar e simplificar a análise desses resultados.

O algoritmo produz gráficos que mostram a frequência com que o *OpenPose* determinou a posição de cada articulação em pixels. O valor médio desses pixels é usado como um ponto de referência, considerado zero. Dessa maneira, é possível avaliar a variação em pixels que o *OpenPose* apresenta ao estimar a posição de um ponto que, em teoria, é estático. Isso auxilia na determinação de uma margem de erro.

Foram gerados gráficos específicos para seis articulações chave para os exercícios em estudo: quadril esquerdo, joelho esquerdo, calcanhar esquerdo (exibido abaixo na figura 1), quadril direito, joelho direito e calcanhar direito. Além destes, gráficos adicionais foram produzidos para mostrar a variação nos ângulos calculados a partir da triangulação dessas articulações:

especificamente os ângulos formados entre quadril, joelho e calcanhar.

Esta análise está em andamento e visa garantir que o *OpenPose* forneça resultados clinicamente válidos, estabelecendo um padrão de confiabilidade no contexto da fisioterapia.

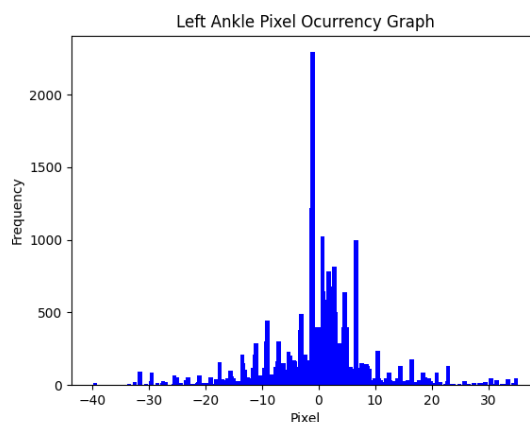


Figura 1 – Gráfico da ocorrência dos pixels no ponto inferido pelo *OpenPose* no tornozelo esquerdo

### 5. Referências

- [1] J. P. R. Caicedo et. al., IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(4):1249–1259, 2014.
- [2] W. Zhang et. al., ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing, pages 5706–5710, 2019.
- [3] M. Saric et. al., 2019 4th International Conference on Smart and Sustainable Technologies, pages 1–4, 2019.
- [4] J. P. Queralt et. al., 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), pages 601–604, 2019.
- [5] H. Chan et. al., 2012 International Symposium on Computer, Consumer and Control, pages 585–588, 2012.
- [6] N. Norouzi-Gheidari et. al., 2013 International Conference on Virtual Rehabilitation (ICVR), pages 220–221, 2013.
- [7] Q. Sanders et. al., IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(6):1363–1372, 2020.
- [8] J. Andreas et. al., ArXiv, abs/1511.02799, 2015.
- [9] J. Andreas et. al., HLT-NAACL, 2016.
- [10] J. A. Francisco et. al., IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, pp. 2174–2183, 2023.

### Agradecimentos

À FAPESP pelo auxílio na aquisição de equipamentos e ao Centro Universitário FEI pelo empréstimo de equipamentos e apoio.

<sup>1</sup> Aluno de IC da FAPESP. Projeto com vigência de 03/2023 a 02/2024.