

ISSN: 2319-0124

O uso de Redes Neurais na Estimativa de Pose e Comparação de Movimentos

Nicolas E. PAIVA¹; Thiago V. G. VALIM²; Alexandre F. SILVA³

RESUMO

Este trabalho visa capturar o movimento executado por um ser humano (usando uma webcam), como por exemplo, de uma pessoa realizando atividade física para o fortalecimento dos bíceps, e compará-la com o mesmo movimento executado por um profissional, de modo que o sistema apresente o quanto os movimentos são diferentes. O sistema será desenvolvido em Python e utilizando redes neurais para o processo de análise e comparação dos movimentos obtidos.

Palavras-chave:

Redes neurais; Python; captura de movimento.

1. INTRODUÇÃO

Analisar a postura humana é uma tarefa desafiadora na visão computacional (Wang et al, 2021). Este processo contém uma grande variedade de subtarefas, como por exemplo, estimativa de pose 3D, rastreamento da pose, etc. (Lim et al, 2011). Métodos computacionais que analisam a postura humana têm sido comumente usados para as mais diversas finalidades, como por exemplo, na análise dos riscos de distúrbios musculoesqueléticos (Kim et al, 2021).

Este trabalho visa capturar o movimento executado por um ser humano (usando uma webcam), como por exemplo, de uma pessoa executando uma atividade física de agachamento, e compará-la com o mesmo movimento executado por um profissional. A ideia é verificar o quanto os movimentos são parecidos e permitir que a pessoa ajuste seus movimentos para que fiquem o mais próximo daquele executado pelo profissional. Para isso, será implementado um algoritmo desenvolvido em Python que captura os movimentos de uma pessoa, a partir de uma webcam, e de uma rede neural para analisar e comparar os movimentos obtidos. Segundo, será criada uma base de dados de movimentos realizados por profissionais da saúde, como por exemplo, o exercício de

¹Bolsista curso integrado EDITAL N°54/2021, IFSULDEMINAS – *Campus* Pouso Alegre. E-mail: nicolas.eduardo@alunos.ifsuldeminas.edu.br

²Bolsista curso integrado EDITAL N°54/2021, IFSULDEMINAS – *Campus* Pouso Alegre. E-mail: thiago.valim@alunos.ifsuldeminas.edu.br

³Orientador, IFSULDEMINAS – *Campus* Pouso Alegre. E-mail: alexandre.silva@ifsuldeminas.edu.br

bíceps, para servir como treinamento da rede neural, e de referência para testar os movimentos capturados pela webcam. Com isso, a comparação em tempo real dos movimentos pode permitir atividades físicas sem um personal trainer profissional por perto.

2. FUNDAMENTAÇÃO TEÓRICA

O estudo do movimento do corpo humano se chama cinesiologia sendo extremamente útil para procedimentos médicos, ou para a análise e aprimoramento do desempenho em atividades físicas (Vyas, 2019). Analisar a pose humana visa determinar a localização de pontos (chaves) no corpo humano (Zheng et al, 2021), e desempenha um papel muito importante nas tarefas da visão computacional (Mao et al, 2021). A estimativa da pose humana tem uma variedade de recursos de aplicações multimídias no mundo real, como por exemplo, na realidade virtual, e na interação com jogos, e podem ser implementadas por meio de técnicas de aprendizado de máquina (Vyas, 2019).

Wang et al (2021) propuseram um sistema de benchmark de pose robusto para avaliar o desempenho de modelos que estimam a pose a partir de imagens corrompidas. Mao et al (2021) propuseram uma estrutura de estimativa de pose que melhora o desempenho da estimativa de pose baseada em métodos de regressão. Assim, o sistema pode capturar os pontos entre as articulações do corpo, de uma maneira mais eficiente.

O objetivo deste trabalho é implementar um sistema capaz de analisar os movimentos de uma pessoa, durante a realização de seu treino, para que possa evitar erros de execução de seus movimentos, assim poderá corrigir e fazer o treino de forma mais eficaz e precisa.

3. MATERIAL E MÉTODOS

Este trabalho visa a implementação de um algoritmo desenvolvido em Python que captura os movimentos de uma pessoa, a partir de uma webcam. Em uma segunda etapa, a ser desenvolvida posteriormente, será utilizada rede neural para analisar e comparar os movimentos obtidos, visando auxiliar nos movimentos realizados por pessoas em tratamento, ou durante suas atividades físicas, com ou sem a supervisão de um profissional.

Para a implementação foi usada a linguagem de programação Python e as bibliotecas OPENCV para o processamento de imagens, MediaPipe para a detecção do movimento, e SYS utilizada para o tratamento e detecção dos movimentos.

Um sistema comum de captura de movimentos consiste de uma webcam e um hardware para realizar a extração das características, para que haja o processamento em tempo real. O trabalho está dividido em 3 etapas: (1) implementação do algoritmo para capturar os movimentos a partir da webcam, implementado e apresentado neste trabalho; (2) estudo e teste de redes neurais e métodos

de captura de movimentos (que será implementado em uma próxima etapa); e, (3) análise dos resultados.

4. RESULTADOS E DISCUSSÕES

Os resultados obtidos até o momento têm sido satisfatórios quanto à captura dos movimentos corporais e a comparação com demais movimentos. Contudo, devido a problemas com o algoritmo de detecção de pose, na qual o sistema detectava como “igual” uma pose que não estava visível e a identificação das conexões (pontos detectados) em cores diferenciadas para facilitar a visualização por parte do usuário.

Sendo assim, devido às problemáticas citadas, foi possível avançar apenas na primeira etapa, que foi capturar o movimento corporal obtido pela câmera e, então, obter características do corpo humano. Em outras palavras, o sistema atribui 32 pontos sendo 3 valores distintos: eixo X, eixo Y e a visibilidade do ponto. Esses pontos são utilizados para comparar a imagem obtida com a imagem armazenada no banco de dados.

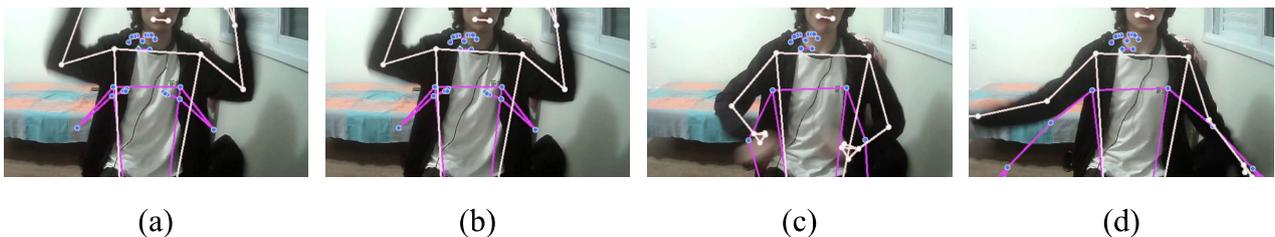


Figura 1 - Movimentos capturados pela câmera (linha branca) comparado com os movimentos armazenados (linha lilás).

A Figura 1 de (a) a (d) apresenta a captura da imagem base (linha lilás) sobreposta à do usuário (linha branca) que é armazenada durante a captura do vídeo. Para realizar o alinhamento das linhas (temporais) e do cálculo da distância entre os pontos, usamos o algoritmo DTW (sigla para *Dynamic Time Warping*), como apresentado na Figura 2.

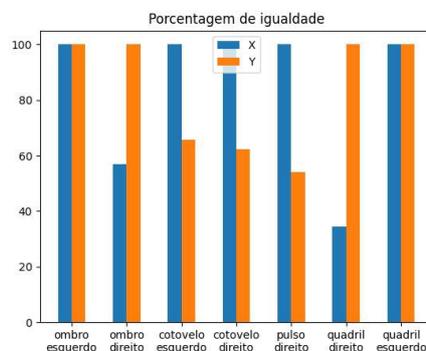


Figura 2 - Resultado do quanto cada ponto capturado (linha branca) é similar ao movimento armazenado (linha lilás)

5. CONCLUSÕES

Um sistema de sistema de captura e comparação de movimentos de exercícios capturados por uma webcam, pode ser utilizado para a avaliação de treinos e posturas. Contudo, a implementação deste sistema que compara e indica a similaridade dos movimentos capturados tem apresentado diversas dificuldades. Neste sentido, avançamos na implementação da comparação dos movimentos, mas ainda não fizemos a etapa de treinamento e teste usando redes neurais, que será iniciado em um novo projeto de pesquisa, como trabalhos futuros. Sendo assim, além da implementação da rede neural, vamos produzir um aplicativo que auxilie os usuários na verificação dos seus movimentos e a similaridade com o movimento treinado (realizado por especialista), a partir de estatísticas e comparação gráfica.

REFERÊNCIAS

- KIM, 2021. KIM, Woojoo et al. Ergonomic postural assessment using a new open-source human pose estimation technology (OpenPose). *International Journal of Industrial Ergonomics*, v. 84, p. 103164, 2021.
- LIM, 2011. LIM, Edwin Choon Wyn et al. Effects of Pilates-based exercises on pain and disability in individuals with persistent nonspecific low back pain: a systematic review with meta-analysis. *Journal of orthopedic & sports physical therapy*, v. 41, n. 2, p. 70-80, 2011.
- MAO, 2021. MAO, Weian et al. Tfpose: Direct human pose estimation with transformers. *arXiv preprint arXiv:2103.15320*, 2021.
- VYAS, 2019. VYAS, Parth. *Pose estimation and action recognition in sports and fitness*. 2019.
- WANG, 2021. WANG, Jiahang et al. When Human Pose Estimation Meets Robustness: Adversarial Algorithms and Benchmarks. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021. p. 11855-11864.
- ZHENG, 2021. ZHENG, Ce et al. 3d human pose estimation with spatial and temporal transformers. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021. p. 11656-11665.