

Desenvolvimento de um Sistema de Monitoramento de Exercícios Fisioterápicos com Auto-Encoder LSTM

Luís Henrique C. Evangelista¹, Carlos Maurício S. Figueiredo¹, Elloá B. Guedes¹

¹ Grupo de Pesquisa em Sistemas Inteligentes
Universidade do Estado do Amazonas (UEA)
Av. Darcy Vargas, 1200 – Manaus – Amazonas

{lhce.snf20, cfigueiredo, ebgcosta}@uea.edu.br

Abstract. *This work aims to contribute to the intelligent monitoring of physiotherapy exercises for telerehabilitation purposes. We propose a Computer Vision solution based on Deep Learning with Auto-Encoders and LSTMs for classifying exercise videos according to the correctness of their execution. Integrating non-invasive pose landmark estimation and anomaly detection techniques, the proposed solution contributes to the monitoring and feedback of remote patients, which can positively collaborate with treatment adherence.*

Resumo. *Este trabalho visa colaborar no monitoramento inteligente de exercícios fisioterápicos para fins de telereabilitação com uma solução de Visão Computacional baseada em Deep Learning com Auto-Encoders LSTM para classificação de vídeos de exercícios quanto à corretude de execução. Integrando técnicas não-invasivas de estimação de pose e de detecção de anomalias, a solução proposta contribui para o acompanhamento e feedback para pacientes remotos, o que pode colaborar positivamente na adesão ao tratamento.*

1. Introdução

De acordo com a Organização Mundial da Saúde (OMS), uma a cada três pessoas vive em uma condição de saúde que se beneficiaria de reabilitação [OMS 2024], e espera-se que a demanda por esse serviço aumente nos próximos anos. Embora haja progressos significativos na área da Fisioterapia, continuam a existir desafios na avaliação do desempenho dos pacientes durante suas atividades fora do ambiente clínico. Adicionalmente, muitos pacientes podem sentir-se indispostos ou incapazes de frequentar regularmente o consultório [Boukhenoufa et al. 2022], o que pode afetar diretamente sua adesão ao tratamento [Subtil et al. 2011]. Há que se mencionar ainda a Pandemia do COVID-19, a qual também ampliou a demanda por serviços de telereabilitação, passando também a abranger um maior público-alvo [Prvu Bettger and Resnik 2020].

Com o desenvolvimento tecnológico, pesquisadores têm proposto novas soluções para auxiliar fisioterapeutas no monitoramento e avaliação clínicos de seus pacientes. Um exemplo de solução neste âmbito inclui o desenvolvimento de diversos tipos de sensores, tais como as Unidades de Medição Inercial (UMI) que medem e registram movimentos tridimensionais [Boukhenoufa et al. 2022]. No âmbito da Visão Computacional (VC), a estimativa de pose humana sem marcadores emerge como uma abordagem promissora no referido contexto [Hellsten et al. 2021]. A estimação de pose humana refere-se ao processo de determinar a posição e a pose (ou seja, a postura ou configuração corporal) de uma pessoa em um espaço tridimensional. Isso envolve a identificação das articulações e segmentos do corpo humano bem como a estimativa de suas posições relativas em relação

a um sistema de coordenadas de referência [Sigal 2021]. A estimação de pose com VC não requer a aplicação de marcadores ou sensores externos no sujeito, consistindo apenas na captura de dados de movimento por meio de imagens. Tais características podem viabilizar uma análise detalhada e supervisão remota de exercícios fisioterapêuticos, mitigando desafios práticos enfrentados tanto por profissionais quanto pelos pacientes.

Com vistas a contribuir com o corpo de soluções para a telereabilitação fisioterapêutica, bem como apoiar profissionais dessa área no acompanhamento de seus pacientes, este trabalho apresenta uma proposta para o monitoramento remoto e inteligente de exercícios de reabilitação. Como estudo de caso, foi escolhido o movimento de levantamento de braço (*arm raise*, em inglês) comumente utilizado em programas de fisioterapia para fortalecer os músculos do ombro, braço e parte superior das costas, melhorando a amplitude de movimento nessas articulações. Este exercício é frequentemente prescrito para pessoas que estão se recuperando de lesões no ombro, como entorses, bursites, tendinites, ou após cirurgias ortopédicas.

A solução proposta visa monitorar de forma objetiva e precisa a execução dos referidos exercícios por meio da análise de imagens capturadas durante sua execução, identificando incorreções com uma arquitetura de *Deep Learning* (DL) baseada em um Auto-encoder LSTM (do inglês, *Long Short-Term Memory*). O modelo proposto foi treinado em uma base de dados de exercícios executados corretamente por diferentes sujeitos. Para isso, pontos do corpo humano foram capturados por meio da técnica de *Pose Landmark*, permitindo a identificação e análise entre os ângulos corporais existentes. Assim, essa abordagem sistemática e baseada em dados provê um diagnóstico preciso e personalizado, otimizando estratégias de tratamento e métodos de ação para o entendimento das dificuldades de cada paciente.

Para apresentar o que se propõe, o presente artigo está organizado como segue. Os trabalhos relacionados são apresentados na Seção 2. A Seção 3 apresenta detalhes da solução proposta, como informações sobre dados experimentais, modelo e avaliação. Em seguida, a Seção 4 mostra os resultados obtidos e, por fim, considerações finais e trabalhos futuros encontram-se dispostos na Seção 5.

2. Trabalhos Relacionados

Resultados do uso e aplicação de técnicas de VC para análise de movimentos de pacientes de fisioterapia, bem como para supervisão remota da execução de exercícios de reabilitação, já foram reportados na literatura. Hellsten *et al.* [2021] discutiram a importância e o potencial de sistemas de análise de movimento humano sem o uso de marcadores, enfatizando a aplicação dessas tecnologias na reabilitação. O artigo destaca como a pandemia do COVID-19 ampliou a demanda por serviços de telereabilitação. Apesar dos avanços, o desafio permanece em desenvolver um sistema eficaz de estimação de pose em 3D com apenas uma câmera que consiga identificar com precisão as coordenadas de todas as juntas corporais, permitindo análises de movimento avançadas.

Dorado *et al.* [2019] descreve um sistema baseado no Microsoft Kinect para a reabilitação em casa de pacientes com artrite reumatoide. Este sistema propõe substituir a supervisão direta do fisioterapeuta, acompanhando a execução dos movimentos, além de coletar e gerenciar informações sobre o histórico das sessões de exercícios. A facilidade de uso da interface é um destaque, considerando o público-alvo principal. Há que se mencionar, entretanto, a necessidade de aquisição de um *hardware* específico para viabilizar sua utilização.

Latreche *et al.* [2023], por sua vez, apresentam um estudo sobre a confiabilidade e validade de um sistema de medição para reabilitação baseado no *MediaPipe* [Lugaresi et al. 2019], aproveitando-se da capacidade deste algoritmo em prever locais de marco de pose para aplicar suavização com o objetivo de obter resultados mais precisos. Este trabalho buscou superar as limitações de ferramentas tradicionais e até mesmo de dispositivos mais caros. Para tanto, contou com a participação de 50 voluntários e métodos tradicionais de VC para aumentar a eficiência e eficácia da solução proposta.

Este conjunto de trabalhos culminou na proposta de soluções que se concentraram não apenas na eficácia da recuperação, mas também na segurança e comodidade dos pacientes. A abordagem aqui proposta visa seguir estes preceitos, mas também contornando a necessidade de aquisição de hardware específico e com flexibilidade para adaptar-se a diferentes sujeitos sem interferência humana ou reajuste de parâmetros, o que diminui o ônus computacional e permite sua utilização em cenários que há restrição no contato entre profissional de saúde e paciente, tornando-a mais flexível perante contextos realistas. Ademais, o uso de técnicas do estado da arte de VC e DL corroboram para uma reabilitação mais acessível e eficiente, avançando o estado da arte não só em tecnologias para a área de Saúde, mas também oferecendo um caminho promissor para melhorar a qualidade de vida de diversas pessoas.

3. Solução Proposta

Guiada pela crescente popularização de ferramentas envolvendo VC, DL e aplicações emergentes envolvendo estimação de pose humana sem marcadores, uma proposta de solução foi concebida no escopo deste trabalho, a qual é descrita na Figura 1. A solução proposta visa monitorar pacientes durante sessões de fisioterapia, bem como identificar falhas na execução dos exercícios realizados. Para tanto, necessita capturar pontos do corpo humano por meio da técnica de *Pose Landmark* e analisar ângulos formados por suas conexões durante a realização de movimentos. Além disso, contemplou a implementação e treino de uma arquitetura de DL própria baseada em um Auto-Encoder LSTM capaz de reconhecer padrões de movimentos ideais e anômalos, tal como exemplificado na Figura 2.



Figura 1. Arquitetura simplificada do sistema de análise de movimentos em exercícios fisioterápicos.

A ferramenta *MediaPipe* foi adotada no escopo deste trabalho para estimação de pose não invasiva a partir de *frames* bidimensionais [Google 2023]. As características coletadas pela mesma foram cruciais para o delineamento do conjunto de dados de treinamento dos modelos de DL, os quais são descritos a seguir.

3.1. Dados Experimentais

Após uma análise criteriosa de diversas bases de dados relacionadas à exercícios fisioterapêuticos, observou-se que muitas delas não se mostravam adequadas para coleta

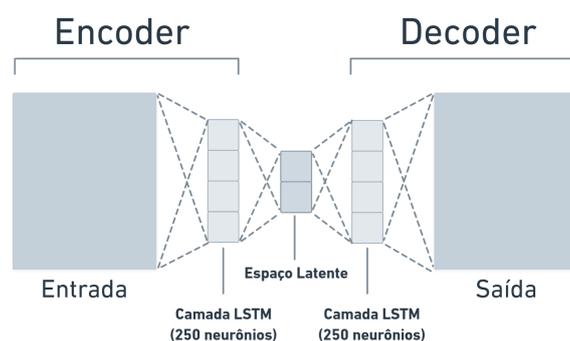


Figura 2. Arquitetura própria do Auto-Encoder LSTM projetada.

de dados de estimativa de pose, não possuíam rótulos sobre o tipo ou corretude da execução de exercícios ou eram diminutas no tocante aos exemplos, o que inviabilizava soluções generalistas, tal como o *dataset* de exercícios fisioterapêuticos UI-PRMD [Vakanski et al. 2018].

Após uma triagem meticulosa, identificou-se que a base de dados sintética InfiniteRep [ToInfinityAI 2023], era viável para os fins almeçados no escopo deste trabalho, pois possuía uma extensa gama de exercícios fisioterápicos, incluindo mais de 1.000 vídeos de diferentes atividades, em que cada uma delas é representada por 100 amostras. Os dados apresentam modelos 3D de indivíduos com variações em características físicas, ritmo e amplitude de movimento, ambientados em diferentes contextos, condições de iluminação e ângulos de câmera. Os vídeos utilizados na etapa de treinamento do modelo tinham a taxa de quadros igual a 48 FPS (do inglês, *Frames Per Second*, ou Quadros Por Segundo). Exemplos de *frames* extraídos da base de dados estão demonstrados na Figura 3.

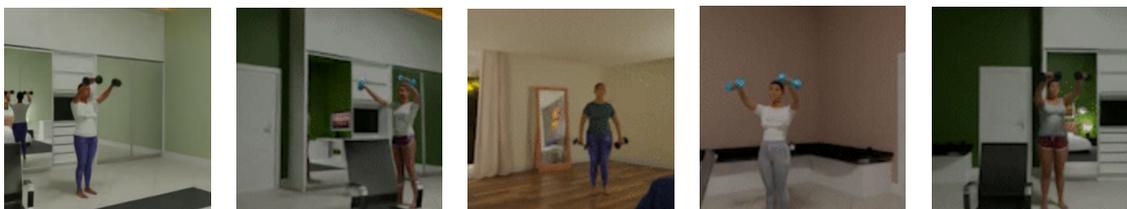


Figura 3. Exemplo de *frames* extraídos do *dataset* InfiniteRep mostrando pessoas realizando o exercício de elevação de braços (*armraise*).

A escolha do InfiniteRep deve-se não apenas a sua diversidade e abrangência, mas também à compatibilidade com as funcionalidades do *MediaPipe*. Isso permite a captura detalhada de uma ampla gama de exercícios fisioterapêuticos e fornece informações precisas sobre a posição e o ângulo das partes do corpo, essenciais para a análise de movimento efetiva.

3.2. Pré-processamento com *MediaPipe* para Extração de *Pose Landmarks*

Com o objetivo de abordar os movimentos estudados de maneira detalhada e precisa, foram selecionados 12 ângulos específicos do corpo humano para análise dentre os 33 rastreados pelo *Pose Landmark* demonstrados na Figura 4. Eles foram registrados em arquivos CSV (do inglês, *Comma Separated Values*) correspondentes aos vídeos da base de dados escolhida, em que cada linha representa um quadro específico do vídeo e contém as informações angulares das partes do corpo selecionadas. Essa estratégia permitiu um

foco detalhado em áreas cruciais para a análise do movimento, facilitando a interpretação dos dados pelo modelo de DL.

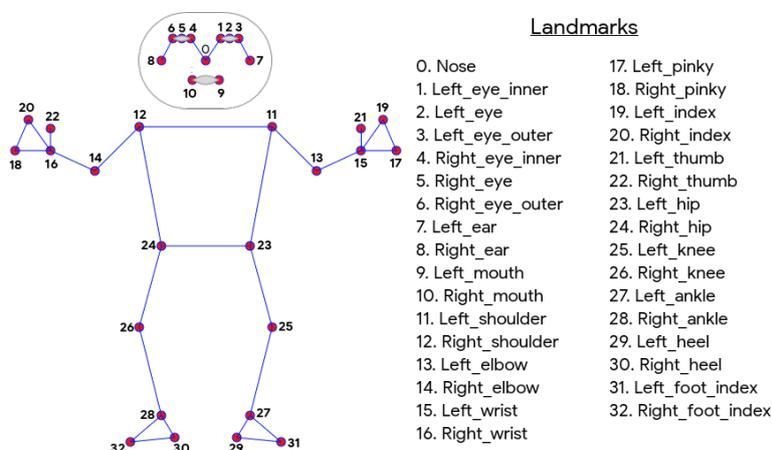


Figura 4. Pontos rastreados pelo Pose Landmark do MediaPipe. Fonte: <https://developers.google.com/ml-kit/vision/pose-detection>

Os ângulos selecionados para análise refletem partes críticas do corpo envolvidas em uma vasta gama de movimentos fisioterápicos e encontram-se detalhados na Tabela 1. A exclusão inicial de outros ângulos, como os dos pulsos, pés e pescoço, foi uma decisão estratégica para simplificar a complexidade do modelo em sua fase inicial, mantendo o foco nas regiões mais impactantes para os movimentos em estudo. No entanto, a inclusão de tais ângulos em futuras iterações do sistema é considerada para enriquecer ainda mais a análise e aumentar a precisão do modelo para tipos de exercícios mais específicos.

Tabela 1. Ângulos selecionados a partir do Pose Landmark.

Ângulo	Especificação
Cotovelo Direito	Medido entre o ombro direito, o cotovelo direito e o pulso direito
Cotovelo Esquerdo	Medido entre o ombro esquerdo, o cotovelo esquerdo e o pulso esquerdo
Ombro Direito	Medido entre o cotovelo direito, o ombro direito e o ombro esquerdo
Ombro Esquerdo	Medido entre o cotovelo esquerdo, o ombro esquerdo e o ombro direito
Joelho Direito	Medido entre o quadril direito, o joelho direito e o tornozelo direito
Joelho Esquerdo	Medido entre o quadril esquerdo, o joelho esquerdo e o tornozelo esquerdo
Quadril Direito	Medido entre o joelho direito, o quadril direito e o quadril esquerdo
Quadril Esquerdo	Medido entre o joelho esquerdo, o quadril esquerdo e o quadril direito
Axila Direita	Medido entre o ombro direito, o quadril direito e o cotovelo direito
Axila Esquerda	Medido entre o ombro esquerdo, o quadril esquerdo e o cotovelo esquerdo
Coxa Direita	Medido entre o quadril direito, o quadril esquerdo e o joelho direito
Coxa Esquerda	Medido entre o quadril direito, o quadril esquerdo e o joelho esquerdo

Após a coleta e organização dos dados, foi implementada a técnica de Aumento Artificial de Dados (AAD, do inglês, *Data Augmentation*) devido ao número relativamente limitado de vídeos disponíveis do movimento em foco desta pesquisa (levantamento de braço ou *arm raise*). Visando simular variações naturais que ocorrem durante a execução dos exercícios, foram adicionados ruídos aleatórios e realizado o deslocamento temporal das informações. Técnicas para tratamento de *outliers* e suavização com Média Móvel também foram aplicadas com o objetivo de melhorar a generalização do modelo.

3.3. Auto-Encoder LSTM para Detecção de Anomalias

Em face das capacidades de aprendizado de sequências temporais complexas das Redes Neurais Recorrentes, optou-se por utilizar a arquitetura LSTM no escopo deste trabalho. A partir disso, a estrutura de DL selecionada para identificação inteligente de padrões corretos e incorretos foi um Auto-Encoder, haja vista sua eficiência em compactar e reconstruir os dados de entrada. Esse arranjo possibilitou a construção de uma arquitetura própria que envolve o uso de um Auto-Encoder com camadas LSTM para codificação (*encoder*) e decodificação (*decoder*), facilitando a identificação de padrões normais de movimento e a detecção de desvios na execução dos exercícios.

O modelo teve as camadas LSTM intercaladas com camadas de *dropout* para mitigar possível *overfitting*, sendo configurado com múltiplas unidades nas camadas de *encoder*, *decoder* e *dropout*, resultando em 16 arquiteturas avaliadas. Adotou-se o uso de *callbacks* para *Early Stopping* e *Model Checkpoint*, visando a otimização da função de perda Erro Quadrático Médio (MSE – do inglês, *Mean Square Error*). A taxa de aprendizado considerada foi de 10^{-4} , bem como o uso do otimizador Adam e da função de ativação ReLU. Optou-se pela aplicação da regularização L2 para evitar a complexidade excessiva das arquiteturas. Os hiperparâmetros do modelo foram ajustados conforme o necessário como o intuito de maximizar a precisão na identificação de padrões de movimento.

3.4. Aplicação de Limiar para Indicação de Movimentos Errados

Após o treinamento do modelo Auto-Encoder LSTM, os Erros de Reconstrução (ER) dos dados de validação foram calculados, o que é fundamental para avaliar a precisão com que o modelo consegue reconstruir os dados de entrada. Para visualizar melhor esta avaliação, um histograma dos ER foi gerado, detalhado na Figura 5. Ele desempenha um papel crucial ao revelar a distribuição dos ER, fornecendo uma base visual e quantitativa para identificar o Limiar de Detecção de Anomalias (LDA), sendo estabelecido no valor que corresponde ao pico da distribuição no histograma, significando o Erro de Reconstrução mais frequente e, por inferência, considerado normal para o conjunto de dados em questão. Definir o LDA desta maneira permite identificar de forma eficaz os movimentos atípicos em análises futuras de novos dados de entrada, como vídeos de exercícios fisioterapêuticos, pois qualquer movimento que gere um ER acima deste limiar será marcado como anômalo.



Figura 5. Histograma de erros de reconstrução e definição do LDA.

4. Resultados e Discussão

Conforme as especificações descritas na seção anterior, os modelos propostos foram implementados e treinados para avaliação do exercício de levantamento de braço, previamente ilustrado na Figura 3, e caracterizado pelo ato de levantar ambos os braços paralelamente até a altura da cabeça. Levando em consideração a preparação de dados realizada, a técnica de validação cruzada *k-folds*, utilizando de 3 a 5 *folds*, foi empregada para garantir uma avaliação robusta e confiável dos modelos, evitando vieses de flutuação estocástica de inicialização de pesos e de outras escolhas de caráter aleatório [Brink et al. 2017].

Para avaliar a capacidade de cada modelo treinado em identificar movimentos corretos, incorretos, bem como as transições entre eles, foram realizados testes práticos com vídeos gravados especificamente para este fim. Ademais, foi feita a definição do valor do LDA, crucial para posterior identificação de movimentos atípicos durante a análise de novos vídeos de exercícios.

A análise dos dados utilizados em diferentes iterações do treinamento do modelo revelou *insights* valiosos. Durante o treino, a janela de Média Móvel, ajustada para 7 quadros, ajudou a suavizar as flutuações temporais nos dados, melhorando a capacidade das arquiteturas de capturar tendências relevantes de movimento. Observou-se também que o Erro de Reconstrução Médio (ERM) e o Erro de Reconstrução Máximo (ERMax) foram influenciados significativamente pelos ajustes nos hiperparâmetros e pelas estratégias de AAD aplicadas, já que permitiu ao modelo adaptar-se melhor às variações naturais dos dados de entrada. As arquiteturas treinadas e seus respectivos hiperparâmetros encontram-se listadas na Tabela 2. Todas apresentaram um ERM variando entre 0,0065 e 0,0135, com o Auto-Encoder LSTM representado pela Arquitetura 1 apresentando o menor valor.

Tabela 2. Avaliação dos Modelos de Auto-encoder LSTM.

Arquiteturas	ERM	ERM _{máx}	LDA ¹	aug_nl ²	aug_rf ³	n_u ⁴	b.s ⁵	folds ⁶
1	0,0065	0,0150	0,0057	0,05	0,3	250	16	5
2	0,0130	0,0220	0,0125	0,05	0,3	150	48	3
3	0,0065	0,0140	0,0056	0,05	0,3	300	48	3
4	0,0099	0,0201	0,0095	0,05	0,5	300	48	3
5	0,0135	0,0224	0,0135	0,10	0,3	150	16	3
6	0,0136	0,0240	0,0129	0,10	0,3	128	16	3
7	0,0188	0,0317	0,0181	0,10	0,3	64	16	3
8	0,0189	0,0297	0,0188	0,10	0,3	32	16	3
9	0,0191	0,0295	0,0182	0,10	0,3	16	16	3
10	0,0187	0,0318	0,0176	0,10	0,3	8	16	3
11	0,0091	0,0187	0,0085	0,05	0,3	256	16	3
12	0,0134	0,0234	0,0131	0,05	0,3	128	16	3
13	0,0088	0,0179	0,0081	0,05	0,5	256	16	3
14	0,0110	0,0199	0,0100	0,05	0,8	128	16	3
15	0,0092	0,0174	0,0088	0,05	0,5	256	16	5
16	0,0119	0,0205	0,0110	0,05	0,5	128	16	5

¹ Limiar de Detecção de Anomalias.

² Nível de Ruído no AAD.

³ Fator de Redução no AAD.

⁴ Número de Unidades na LSTM.

⁵ Tamanho do lote (*batch_size*).

⁶ Número de *folds* utilizados na Validação Cruzada.

O valor do LDA, ajustado com base no desempenho do modelo durante a fase de

validação e análise dos ER, mostrou-se uma métrica crucial na identificação precisa de movimentos atípicos, sendo definido com base no pico da distribuição dos ER, cujo valor atribuído foi 0,018. Foi observado que a configuração de AAD, incluindo o nível de ruído e o fator de redução, foram essenciais para gerar variações nos dados de treinamento que, embora não idênticas, mantinham-se alinhadas aos padrões originais. Essa abordagem foi fundamental para enriquecer o conjunto de dados com exemplos variados, aumentando a capacidade do modelo de generalizar para novos dados sem perder a consistência com os movimentos esperados, contribuindo assim para a robustez do modelo.

Os parâmetros de treinamento, em especial o tamanho do lote e o número de épocas, foram definidos considerando a eficiência do treinamento e o desempenho do modelo. Embora o número de épocas tenha sido estabelecido inicialmente em 600 para permitir uma ampla margem de aprendizado por meio do ajuste de parâmetros livres, a implementação da técnica de regularização *Early Stopping*, com uma paciência configurada de maneira estratégica, tornou o número total de épocas menos relevante. O modelo com o melhor desempenho foi configurado com um tamanho de lote de 48, destacando a importância da configuração adequada da paciência para maximizar a eficiência e eficácia no processo de aprendizado.

Em síntese, os resultados obtidos destacam a eficácia das técnicas de aumento de dados e dos ajustes cuidadosos dos hiperparâmetros na melhoria da precisão e na robustez do modelo de Auto-Encoder LSTM para a análise de movimentos em exercícios fisioterápicos. Uma representação visual da detecção dos pontos-chave e da classificação atribuída aos movimentos realizados pela solução proposta encontra-se ilustrada na Figura 6. Neste cenário ilustrativo, o termo *loss* refere-se ao desempenho do modelo após a avaliação do último segmento do vídeo. Caso a *loss* exceda o limite pré-definido, isto é, o ERMáx para o modelo específico, o resultado é classificado como insatisfatório, isto é, o movimento realizado não está em conformidade com o esperado.

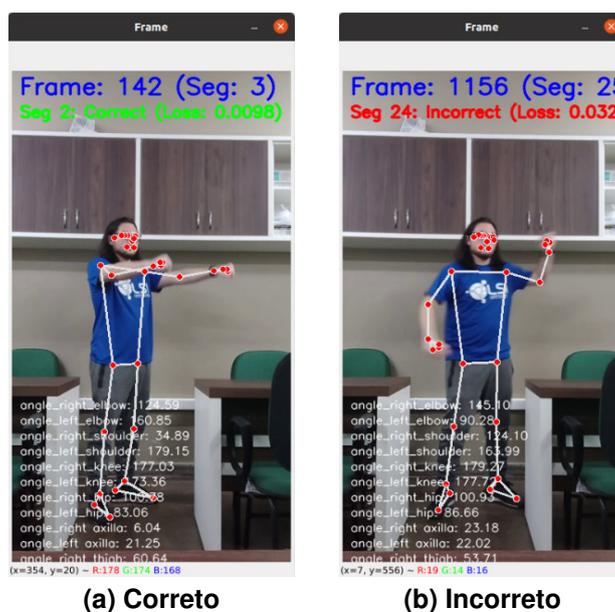


Figura 6. Classificação de movimentos pelo Auto-Encoder LSTM treinado.

A investigação envolveu a análise de segmentos de 24 quadros do vídeo em questão, os quais foram rotulados como corretos ou incorretos com base na proporção

de movimentos executados adequadamente durante esse intervalo. Notavelmente, um segmento foi considerado correto mesmo que apenas o final apresentasse um movimento incorreto. Sob esses critérios, os modelos foram rigorosamente avaliados, destacando-se aquele exposto na primeira linha da Tabela 2, que incorpora uma camada de codificação e uma de decodificação com 250 neurônios LSTM em cada. Esse modelo obteve uma acurácia igual a 68%. A matriz de confusão, mostrada na Tabela 3, revela que a principal vulnerabilidade da arquitetura reside na identificação de movimentos corretos, evidenciado pelo fato de que, em 8 dos 17 movimentos regulares testados, incorreram em classificações do tipo Falso Negativo.

Tabela 3. Matriz de Confusão dos experimentos.

Previsto	Real	
	Correto	Incorreto
Correto	9	8
Incorreto	0	8

Como resultado do desenvolvimento deste trabalho, além da arquitetura própria apresentada, tem-se um protótipo de ferramenta cuja a interface do usuário foi desenvolvida em Flask, permitindo aos usuários carregar vídeos de exercícios fisioterápicos para análise. Utiliza-se do OpenCV e *MediaPipe* para processamento de vídeo e análise de pose, dividindo os vídeos em segmentos a serem avaliados pelo modelo. Os resultados são armazenados em arquivos no formato CSV, facilitando a avaliação da execução dos exercícios e a detecção de anomalias. A Figura 7 demonstra uma captura de tela da interface do usuário.

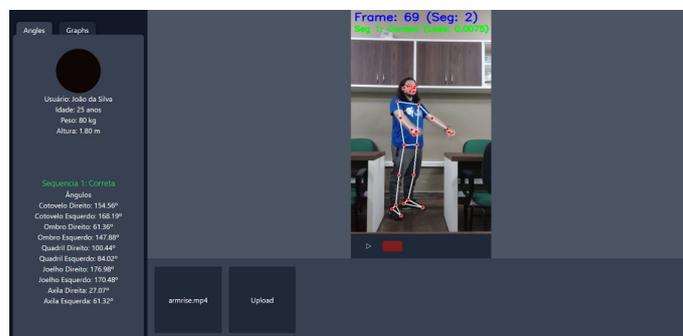


Figura 7. Visualização da interface do usuário, mostrando a análise de movimento em tempo real e *feedback* visual.

5. Considerações Finais

Os resultados deste trabalho destacam o potencial da aplicação de Auto-Encoders LSTM na análise de exercícios fisioterápicos quanto à sua corretude de execução. Para tanto, considerou o movimento de levantamento de braço e diversos vídeos com sujeitos em diferentes condições de captura dos dados (posição, iluminação, etc.), exemplos estes obtidos a partir de uma base de dados disponível na literatura. Diversas arquiteturas foram propostas, treinadas e avaliadas segundo a estratégia *k-folds*, o que permitiu propor uma solução inteligente e não-invasiva que visa colaborar na telereabilitação de pacientes.

A solução proposta não se propõe a substituir especialistas humanos na supervisão de exercícios fisioterápicos. Para tanto, mais pesquisas e experimentos controlados precisam ser realizados. Em trabalhos futuros, almeja-se incluir mais variáveis associadas à

posição de punho, mãos e pés com vistas a enriquecer o *feedback* ao usuário sobre como contornar a incorreção na execução de um determinado exercício, bem como ampliar a gama de exercícios fisioterápicos considerados, o que pode contribuir para uma solução mais personalizada e de mais ampla adoção.

Referências

- Boukhennoufa, I., Zhai, X., Utti, V., Jackson, J., and McDonald-Maier, K. D. (2022). Wearable sensors and machine learning in post-stroke rehabilitation assessment: A systematic review. *Biomedical Signal Processing and Control*, 71:103197.
- Brink, H., Richards, J. W., and Fetherolf, M. (2017). *Real-World Machine Learning*. Manning Publications, Estados Unidos.
- Dorado, J., del Toro, X., Santofimia, M. J., Parreño, A., Cantarero, R., Rubio, A., and Lopez, J. C. (2019). A computer-vision-based system for at-home rheumatoid arthritis rehabilitation. *International Journal of Distributed Sensor Networks*, 15(9):1550147719875649.
- Google (2023). Pose landmarker solution. Acesso em 24 de junho de 2023: https://developers.google.com/mediapipe/solutions/vision/pose_landmarker.
- Hellsten, T., Karlsson, J., Shamsuzzaman, M., and Pulkkis, G. (2021). The potential of computer vision-based marker-less human motion analysis for rehabilitation. *Rehabilitation Process and Outcome*, 10:11795727211022330.
- Latreche, A., Kelaiaia, R., Chemori, A., and Kerboua, A. (2023). Reliability and validity analysis of mediapipe-based measurement system for some human rehabilitation motions. *Measurement*, 214:112826.
- Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., et al. (2019). Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*. Acesso em 26 de junho de 2023: <https://arxiv.org/abs/1906.08172>.
- OMS (2024). Rehabilitation. Disponível em https://www.who.int/health-topics/rehabilitation#tab=tab_1. Acesso em 24 de maio de 2024.
- Prvu Bettger, J. and Resnik, L. J. (2020). Telerehabilitation in the age of COVID-19: An opportunity for learning health system research. *Phys. Ther.*, 100(11):1913–1916.
- Sigal, L. (2021). *Human Pose Estimation*, pages 573–592. Springer International Publishing, Cham.
- Subtil, M. M. L., Goes, D. C., Gomes, T. C., and Souza, M. L. d. (2011). O relacionamento interpessoal e a adesão na fisioterapia. *Fisioterapia em Movimento*, 24(4):745–753.
- ToInfinityAI (2023). Infiniterrep dataset. <https://github.com/toinfinityai/InfiniteRep>.
- Vakanski, A., Jun, H.-p., Paul, D., and Baker, R. (2018). A data set of human body movements for physical rehabilitation exercises. *Data*, 3(1):2.