



UNIVERSIDADE
ESTADUAL DE LONDRINA

BRUNO GIOVANINI

**VIABILIDADE DE UMA ABORDAGEM BASEADA EM VÍDEO
UTILIZANDO DEEP LEARNING PARA RASTREAMENTO
AUTOMÁTICO DE JOGADORES DE BASQUETEBOL**

Londrina
2025

BRUNO GIOVANINI

**VIABILIDADE DE UMA ABORDAGEM BASEADA EM VÍDEO
UTILIZANDO DEEP LEARNING PARA RASTREAMENTO
AUTOMÁTICO DE JOGADORES DE BASQUETEBOL**

Dissertação apresentada ao programa de Pós-Graduação associado em Educação Física – UEL/UEM, como requisito para à obtenção do título de Mestre em Educação Física.

Orientador: Prof. Dr. Felipe Arruda Moura.

Londrina
2025

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

G512v Giovanini, Bruno.
VIABILIDADE DE UMA ABORDAGEM BASEADA EM VÍDEO UTILIZANDO DEEP LEARNING PARA RASTREAMENTO AUTOMÁTICO DE JOGADORES DE BASQUETEBOL / Bruno Giovanini. - Londrina, 2025.
60 f. : il.

Orientador: Felipe Arruda Moura.
Dissertação (Mestrado em Educação Física) - Universidade Estadual de Londrina, Centro de Educação Física e Esportes, Programa de Pós-Graduação em Educação Física, 2025.
Inclui bibliografia.

1. visão computacional - Tese. 2. cinemática linear - Tese. 3. basquetebol - Tese. I. Moura, Felipe Arruda. II. Universidade Estadual de Londrina. Centro de Educação Física e Esportes. Programa de Pós-Graduação em Educação Física. III. Título.

CDU 796

BRUNO GIOVANINI

**VIABILIDADE DE UMA ABORDAGEM BASEADA EM VÍDEO
UTILIZANDO DEEP LEARNING PARA RASTREAMENTO
AUTOMÁTICO DE JOGADORES DE BASQUETEBOL**

Dissertação apresentada ao programa de Pós-Graduação associado em Educação Física – UEL/UEM, como requisito para à obtenção do título de Mestre em Educação Física.

BANCA EXAMINADORA

Orientador: Prof. Dr. Felipe Arruda Moura
Universidade Estadual de Londrina - UEL

Prof. Dr. Bruno Luiz de Souza Bedo
Universidade de São Paulo - USP

Prof. Dr. Hécio Rossi Gonçalves
Universidade Estadual de Londrina - UEL

Londrina, 05 de julho de 2025.

GIOVANINI, Bruno. **Viabilidade de uma abordagem baseada em vídeo utilizando *deep learning* para rastreamento automático de jogadores de basquetebol.** Dissertação (Mestrado em Educação Física) – Universidade Estadual de Londrina, Londrina, 2025.

RESUMO

Avanços na área de visão computacional e inteligência artificial tornaram possível o rastreamento por vídeo utilizando algoritmos de rastreamento de múltiplos objetos. Um desses métodos é o DeepLabCut (DLC), uma ferramenta emergente nas ciências do movimento, que expandiu os escopos de pesquisa por permitir análises quantitativas do movimento em configurações mais ecológicas sem utilizar de equipamentos com alto custo financeiro. Entretanto, até o momento, não existem investigações testando a viabilidade de aplicação de um sistema de rastreamento automático que seja válido e apresente boa qualidade de rastreamento. Portanto, a presente dissertação teve dois objetivos: 1) avaliar a validade concorrente do DLC para estimar a posição e velocidade de jogadores durante tarefas estáticas e dinâmicas e 2) avaliar a qualidade do rastreamento automático de jogadores de basquetebol utilizando o DLC durante um jogo oficial de basquetebol. Quanto ao primeiro objetivo, dez jogadores jovens de basquetebol foram submetidos a tarefas estáticas e dinâmicas (caminhar, trotar, correr e *lane-drill*). O rastreamento utilizando o DLC foi comparado a um sistema optoeletrônico. A validade foi analisada por meio de erros absolutos e relativos, análise gráfica de Bland-Altman, coeficiente de correlação intraclassa (ICC), erro quadrático médio (RMSE) e mapeamento paramétrico estatístico (SPM1d) para as séries temporais de velocidade. De maneira geral, os valores de ICC mostraram boa a excelente confiabilidade entre sistemas. Os erros foram maiores à medida que a velocidade proposta das tarefas aumentou – erros relativos entre 1,00 e 2,11% para distância percorrida e RMSE entre 0,12 e 0,41 m·s⁻¹ para velocidade. A SPM1d indicou diferenças significativas ($p < 0,05$) nas curvas de velocidade dos dois sistemas durante momentos de mudança de direção de 90 graus em velocidades mais altas. Quanto ao segundo objetivo, foi filmado um jogo profissional de basquetebol masculino da primeira divisão nacional utilizando apenas uma câmera fixada sobre a quadra capturando toda a área de jogo. Foram testadas diferentes abordagens de processamento dos dados do DLC e avaliadas métricas de detecção e associação calculadas a partir do registro de predições corretas, predições incorretas e trocas de identidade. De maneira geral, a abordagem utilizando interpolação do tipo '*rloess*' apresentou a melhor qualidade de rastreamento dentre as abordagens testadas – alcançando uma taxa de trocas de apenas 0,23% e um MOTA de 0,982. O tempo estimado de intervenção manual para correção das trocas em cada quarto de jogo é de aproximadamente 70 minutos. Portanto, a utilização do DLC como base de um sistema de rastreamento automático de jogadores de basquetebol se mostrou altamente viável – com validade e qualidade de rastreamento adequadas, além de uma interface flexível, custos logísticos-operacionais acessíveis para a maior parte dos clubes profissionais de basquetebol e podendo entregar resultados confiáveis sobre as variáveis cinemáticas dos jogadores analisados horas após a partida.

Palavras-chave: cinemática linear, basquetebol, rastreamento, visão computacional.

GIOVANINI, Bruno. **Feasibility of a video-based deep learning approach for automatic tracking of basketball players**. Dissertation (master's degree in physical education) – State University of Londrina, Londrina, 2025.

ABSTRACT

Advances in computer vision and artificial intelligence have made it possible to perform video tracking using multiple-object tracking algorithms. One such method is DeepLabCut (DLC), an emerging tool in movement sciences that expands research possibilities by allowing quantitative analysis of movement in more ecological settings without requiring expensive equipment. However, there have been no investigations testing the feasibility of applying an automatic tracking system that is both valid and capable of delivering high-quality tracking. Therefore, this master's dissertation had two aims: 1) to evaluate the concurrent validity of DLC for estimating the position and velocity of basketball players during static and dynamic tasks, and 2) to assess the quality of automatic tracking of basketball players using DLC during an official basketball game. Regarding the first objective, ten young basketball players performed static and dynamic tasks (walking, jogging, running, and lane-drill). Tracking using DLC was compared to an optoelectronic system. Validity was assessed through absolute and relative errors, Bland-Altman graphical analysis, intraclass correlation coefficient (ICC), root mean square error (RMSE), and statistical parametric mapping (SPM1d) for velocity time series. Overall, ICC values indicated good to excellent reliability between systems. Errors increased as task velocities increased – relative errors between 1.00 and 2.11% for distance traveled, and RMSE between 0.12 and 0.41 m·s⁻¹ for velocity. The SPM1d showed significant differences ($p < 0.05$) in velocity curves between the two systems during 90-degree directional changes at higher velocities. Regarding the second objective, a professional men's basketball game from the national first division was filmed using only one fixed camera positioned over the court capturing the entire playing area. Different DLC data processing approaches were tested, and detection and association metrics were evaluated based on correct predictions, incorrect predictions, and identity switches. Overall, the approach using 'rloess' interpolation presented the best tracking quality among the tested methods — achieving an identity switch rate of just 0.23% and a MOTA of 0.982. The estimated manual intervention time required to correct identity switches per quarter is approximately 70 minutes. Therefore, using DLC as the basis for an automatic basketball player tracking system proved highly feasible —demonstrating adequate validity and tracking quality, along with a flexible interface, affordable logistical and operational costs for most professional basketball clubs, and the capability to deliver reliable kinematic data on analyzed players within hours after the game.

Keywords: linear kinematics, basketball, tracking, computer vision.

SUMÁRIO

1	INTRODUÇÃO GERAL	7
2	OBJETIVOS	12
3	PROCEDIMENTOS METODOLÓGICOS	13
3.1	Participantes e aquisição de imagens.....	13
3.2	Tarefas estáticas e dinâmicas	13
3.3	Calibração da área de interesse	14
3.4	Treinamento da rede neural.....	15
3.5	Captura optoeletrônica	16
3.6	Processamento de dados	17
3.7	Análise estatística	19
3.8	Avaliação da qualidade de rastreamento.....	20
4.	RESULTADOS.....	22
4.1	Estudo 1: rastreamento automático de jogadores de esportes de quadra indoor utilizando uma abordagem de <i>deep learning</i> baseada em vídeo: um estudo de validade concorrente	22
4.2	Estudo 2: qualidade de rastreamento automático de jogadores de basquetebol utilizando o deeplabcut tm durante um jogo oficial	38
5	CONCLUSÃO GERAL.....	51
6	REFERÊNCIAS BIBLIOGRÁFICAS	52
	ANEXOS.....	58

1. INTRODUÇÃO GERAL

Avanços tecnológicos têm transformado a análise de desempenho esportivo nos últimos anos, impulsionados pelo crescente volume de informações disponíveis e pela ampla variedade de suas aplicações práticas (COLOMER et al., 2020; LORD et al., 2020). Nesse sentido, diferentes ferramentas biomecânicas para captura do movimento têm avançado para atender aos desafios da área esportiva: de notações manuais a sistemas inerciais, sistemas de posicionamento global ou local, sistemas optoeletrônicos e sistemas sem marcadores utilizando algoritmos de visão computacional (TORRES-RONDA et al., 2022). Por meio de tais ferramentas, é possível avaliar padrões de movimento de forma menos restrita e abrangente. Entre suas aplicações, mas não limitadas a estas, incluem-se: avaliação do risco de lesão ou processo reabilitativo, avaliação das adaptações ao treinamento, avaliação da qualidade técnica ou da eficiência em determinadas habilidades motoras e rastreamento da movimentação de um ou mais atletas (RICO-GONZÁLEZ; LOS ARCOS; ET AL., 2020; TORRES-RONDA ET AL., 2022). O uso isolado ou a combinação de diferentes ferramentas pode variar de acordo com a aplicação e as demandas necessárias em cada contexto de análise do desempenho esportivo (BUCHHEIT et al., 2014).

Recentemente, sistemas para a captura de movimento automática sem utilizar marcadores (ou sistemas *markerless*) têm chamado a atenção de profissionais e pesquisadores pois: (1) superam as limitações estruturais e logísticas necessárias para sistemas inerciais ou optoeletrônicos (COLYER et al., 2018) e (2) permitem análises mais ecológicas, expandindo seu escopo para ambientes não-controlados fora do laboratório, como o treinamento e a competição (SEETHAPATHI et al., 2019). Estas técnicas são baseadas em vídeo e requerem a estimativa da posição e orientação (pose) de um objeto ao longo de uma sequência de imagens. Por meio da identificação de características do objeto em imagens sucessivas, as alterações na pose do objeto podem ser acompanhadas em função do tempo (MOURA; CAETANO; TORRES, 2025).

Tais técnicas têm sido amplamente aplicadas em estudos de biomecânica e análise de movimento e possuem duas aplicações importantes na área da análise de desempenho esportivo. A primeira refere-se à análise cinemática de tarefas simples e específicas do esporte, seja em ambientes controlados (e.g., laboratório ou

simulações) ou não-controlados (e.g., treinamento e competições reais). Por exemplo, estudos anteriores utilizando estas técnicas têm encontrado resultados promissores analisando a cinemática de diferentes tarefas como caminhar (NEEDHAM et al., 2021), correr (JOHNSON, C. D.; OUTERLEYS; DAVIS, 2022; NEEDHAM et al., 2021), saltar (DRAZAN et al., 2021; NEEDHAM et al., 2021) e movimentos subaquáticos (CRONIN et al., 2019; PAPIC et al., 2021). No contexto de tarefas esportivas específicas, outros autores utilizaram diferentes sistemas *markerless* de captura de movimento no futebol (MONTEIRO et al., 2024; PALUCCI VIEIRA et al., 2022), tênis (TAN; CHOW; KOMAR, 2024) e basquetebol (PANG; ZHANG, 2025). Em sua maioria, tais aplicações tiveram como objetivo analisar o padrão de movimento de uma tarefa específica (e.g., o chute no futebol ou o arremesso no basquetebol) para entender os determinantes do desempenho ou possíveis riscos associados a lesões, por exemplo. É razoável dizer que cada objetivo, em cada esporte, terá desafios exclusivos e que os proponentes de novas tecnologias – bem como seus usuários – devem estar atentos às possíveis soluções.

A segunda aplicação importante que sistemas *markerless* podem ter na área da análise de desempenho esportivo é a de rastreamento de atletas em seu campo de jogo durante o treinamento ou competição. Uma revisão recente de Moura e colaboradores (2025) mostrou como sistemas de rastreamento baseados em vídeo tornaram-se ainda mais promissores como o avanço das redes neurais convolucionais e das abordagens de *deep learning* aceleradas por GPU. Nesse tipo de aplicação, são utilizados algoritmos cuja finalidade é rastrear múltiplos objetos – ou indivíduos – ao longo de um vídeo, comumente referenciados como *multi-object tracking* (MOT) ou *multi-person tracking* (MPT). Seja utilizando técnicas *bottom-up* ou *top-down*, estes algoritmos tentam detectar múltiplos indivíduos em uma cena, atribuir uma identidade a cada um e (re)associar esta identidade ao mesmo indivíduo ao longo do vídeo, mesmo havendo variações na posição do indivíduo e nas condições da imagem (LUITEN et al., 2021; LUO et al., 2021; MATHIS et al., 2020). Entre outros, alguns dos mais utilizados são o CMU-Pose (CAO, Z. et al., 2016), o OpenPose (CAO, Z. et al., 2018), o AlphaPose (FANG et al., 2022) e o YOLO-Pose (MAJI et al., 2022). O quadro abaixo apresenta as definições de termos comumente utilizados na área de visão computacional (Quadro 1).

Quadro 1 – Definição dos termos utilizados na área de visão computacional.

Markerless	Técnica de captura de movimento sem a utilização de marcadores fixados no corpo, na qual apenas câmeras e algoritmos de visão computacional e <i>machine learning</i> são utilizados para estimar, rastrear e analisar os movimentos dos indivíduos em cena.
Inteligência artificial	Campo da computação que estuda a criação de máquinas ou sistemas capazes de simular a inteligência humana, realizando tarefas que normalmente exigiriam capacidades cognitivas humanas (por exemplo, percepção, aprendizado e decisão).
Machine learning	Subcampo da inteligência artificial focado no desenvolvimento de algoritmos que aprendem padrões a partir de dados, ajustando seu desempenho em tarefas específicas conforme expostos a novos exemplos, em vez de seguir somente regras fixas programadas.
Deep learning	Subárea do <i>machine learning</i> que utiliza redes neurais artificiais com múltiplas camadas (redes “profundas”) para aprender representações e padrões complexos nos dados, muitas vezes alcançando níveis de abstração mais elevados em problemas de visão, áudio ou texto.
Bottom-up	Abordagem de processamento em que se identificam primeiro os elementos básicos nos dados para depois agrupá-los em estruturas de nível superior; por exemplo, detectam-se todas as articulações (pontos-chave) de todos os jogadores na imagem e então associa-se esses pontos a cada indivíduo para compor suas poses completas.
Top-down	Abordagem de processamento em que primeiro se detectam as entidades de alto nível para então detalhar seus componentes; no contexto de pose humana, localiza-se cada jogador na cena (por meio de um detector de pessoas) e em seguida estimam-se as posições de suas articulações por meio de um modelo de pose.
Transfer learning	Técnica de aprendizado de máquina na qual um modelo treinado em uma tarefa (ou domínio) de origem é reaproveitado em outra tarefa alvo relacionada, transferindo conhecimento prévio para melhorar o desempenho com menos necessidade de dados ou tempo de treinamento adicionais.

Fonte: Wani e colaboradores (2020).

Uma alternativa a estes algoritmos é o DeepLabCut (DLC), um pacote aberto de ferramentas que utiliza redes neurais profundas que podem aprender características importantes de uma imagem sendo treinada com um pequeno conjunto prévio de

exemplos nos quais o usuário rotula os locais corretos de interesse (LAUER et al., 2022; MATHIS et al., 2020). O DLC é baseado no modelo pré-treinado DeeperCut (INSAFUTDINOV et al., 2016), que é baseado na arquitetura ResNet treinada no ImageNet (RUSSAKOVSKY et al., 2015). Diferentemente de alguns outros algoritmos, que também usam redes neurais treinadas com bancos de dados de milhares de imagens rotuladas e pontos-chave padronizados, o DLC incorpora técnicas de *transfer learning*, permitindo à rede "aprender" novos pontos-chave com relativamente poucas imagens de treinamento (LAUER et al., 2022; MATHIS et al., 2018, 2020). Sendo assim, o DLC utiliza uma abordagem *bottom-up*, na qual os pontos-chave são localizados e as redes neurais treinadas são responsáveis por conectar estes pontos e determinar a pose dos indivíduos (MATHIS et al., 2020). Esse tipo de característica aplicada ao rastreamento de atletas em seu campo de jogo pode permitir uma maior facilidade na detecção simultânea de atletas similares com identidades distintas interagindo em proximidade. Ademais, o DLC destaca-se por alguns fatores: (1) pela facilidade de personalização e adaptação dos pontos-chave; (2) permite o usuário tornar o modelo treinado mais ou menos complexo dependendo dos seus objetivos – podendo tornar seu uso mais eficiente utilizando menos recursos computacionais e (3) tem uma interface de fácil usabilidade desde a rotulação até o treinamento do modelo.

Para que um sistema de rastreamento de atletas seja considerado viável – utilizando o DLC ou qualquer outro algoritmo – é necessário que este sistema cumpra alguns requisitos fundamentais para atestar sua viabilidade no contexto prático. O primeiro deles é a validade – ou seja, o sistema deve ser capaz de medir o que se propõe a medir a respeito das variáveis cinemáticas decorrentes do rastreamento. Para isso, é necessário comparar as estimativas do sistema com as medidas de um sistema padrão-ouro. O segundo requisito fundamental é o de qualidade de rastreamento – o sistema deve ser consistente na detecção e associação dos indivíduos que se propõe a rastrear. Outros aspectos importantes da viabilidade dizem respeito aos custos logísticos, computacionais e humanos necessários para a aplicação prática. Sistemas viáveis buscam minimizar esses custos de modo que o tempo necessário para configuração, aplicação e obtenção dos resultados sejam condizentes com as demandas práticas que se busca atender para profissionais e pesquisadores.

Alguns estudos já buscaram avaliar a validade de sistemas de rastreamento baseado em vídeo utilizando tarefas específicas e um sistema padrão-ouro como referência no taekwondo (BANKS et al., 2024) e no futebol (LINKE; LINK; LAMES, 2018). Esses estudos avaliaram a validade de diferentes sistemas para estimar o deslocamento de atletas realizando tarefas específicas da modalidade em circuitos ou combates simulados e tiveram resultados promissores. No entanto, é importante destacar que os sistemas padrão-ouro para captura de movimento (como optoeletrônicos, por exemplo) apresentam limitações quanto a capacidade de área a ser capturada e a quantidade de atletas simultâneos. Essas e outras limitações podem afastar a validação do contexto prático, onde os desafios mencionados acima acontecem em uma escala maior do que os sistemas de referência podem mensurar. Portanto, embora sistemas de rastreamento sejam considerados válidos nos contextos estudados, os resultados da validação devem ser generalizados com cuidado – sendo necessária a avaliação de qualidade de rastreamento do mesmo sistema em contexto prático.

A avaliação da qualidade de rastreamento no contexto prático testará o sistema frente a desafios exclusivos do contexto esportivo, como: a presença de oclusões frequentes, a variabilidade na iluminação e no ambiente de jogo (e.g. presença dos árbitros ou da torcida), movimentos em alta velocidade e mudanças abruptas de direção (CAO ET AL., 2024; MANAFIFARD; EBADI; ABRISHAMI MOGHADDAM, 2017; RICO-GONZÁLEZ ET AL., 2020b). Embora a qualidade de rastreamento de alguns algoritmos já seja conhecida, esta é tradicionalmente avaliada utilizando grandes bases de dados não específicas do esporte, como multidões e ambientes urbanos movimentados (LEAL-TAIXÉ ET AL., 2015; RUSSAKOVSKY ET AL., 2015), não contemplando os desafios do contexto esportivo. A qualidade de rastreamento, nesses casos, é comumente avaliada por três características principais: erros de detecção, erros de associação e erros de localização (LEICHTER; KRUPKA, 2012). Erros de detecção ocorrem quando o algoritmo detecta atletas que não existem ou falha ao detectar atletas que existem na imagem. Erros de associação são aqueles em que o algoritmo atribui a identidade errada ao atleta na imagem. Erros de localização ocorrem quando a posição das detecções não está espacialmente alinhada com a posição real do atleta. Estudos anteriores que avaliam a qualidade de rastreamento de

algoritmos em contexto prático no basquetebol tiveram resultados promissores (FU et al., 2020; KONG; HUANG; WANG, 2020; WU et al., 2019b; ZHANG et al., 2020), porém a acurácia das medidas cinemáticas destes sistemas ainda é desconhecida.

Diante do exposto acima, esta dissertação de mestrado será dividida em dois estudos independentes que buscarão testar a viabilidade do DLC como um sistema de rastreamento automático de atletas de basquetebol. Para isso, será avaliada (1) a validade concorrente do DLC para estimar a posição e velocidade de jogadores durante tarefas estáticas e dinâmicas e, na sequência, (2) a qualidade do rastreamento automático de jogadores de basquetebol utilizando o DLC durante jogos oficiais. De maneira geral, é esperado que o DLC apresente validade concorrente aceitável para estimar a posição e velocidade dos jogadores em relação ao sistema padrão-ouro utilizado. Além disso, também é esperado que o DLC apresente uma qualidade de rastreamento adequada durante jogos oficiais de basquetebol, mantendo erros de detecção e associação dentro dos estabelecidos como aceitáveis pela literatura (MOURA; CAETANO; TORRES, 2025). Além disso, para atestar sua viabilidade, esperamos implementar tal sistema visando eficiência de custos (logísticos e computacionais) e de tempo para facilitar a adesão desse tipo de sistema por organizações esportivas.

2. OBJETIVOS

Os objetivos gerais desta dissertação são: (1) testar a validade concorrente do DLC para estimar a posição e velocidade de jogadores durante tarefas estáticas e dinâmicas e (2) avaliar a qualidade do rastreamento automático de jogadores de basquetebol utilizando o DLC durante jogos oficiais. Além disso, reportaremos o tempo de processamento de dados, o tempo de intervenção manual e os custos computacionais e logísticos do sistema utilizado para avaliar sua viabilidade.

3. PROCEDIMENTOS METODOLÓGICOS

3.1. PARTICIPANTES E AQUISIÇÃO DE IMAGENS

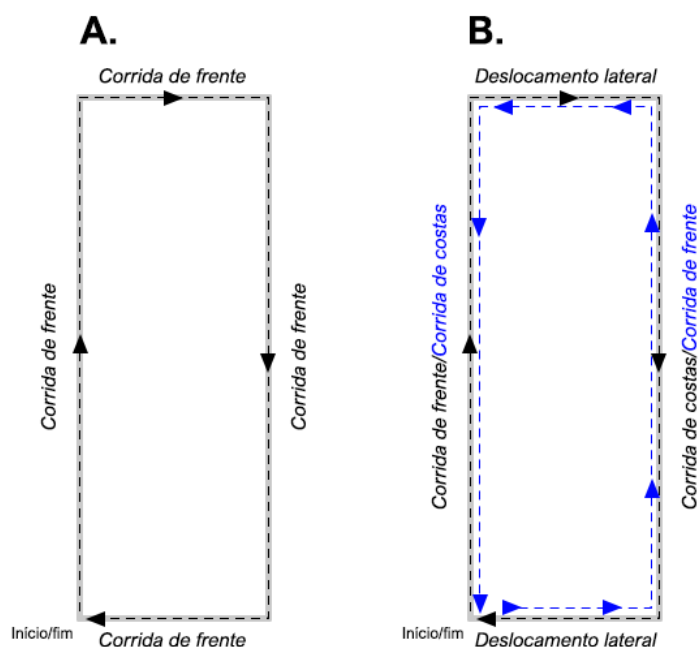
No primeiro estudo, a amostra foi composta, por conveniência, por dez atletas jovens de basquetebol (idade: $16,5 \pm 0,5$ anos; massa: $75,7 \pm 3,6$ Kg; estatura: $181,7 \pm 2,4$ cm) pertencentes às equipes das categorias de base de Londrina – PR. Eles pertenciam à mesma equipe que participava de competições estaduais e treinavam cinco dias por semana. Antes do início da coleta de dados, os participantes e seus responsáveis legais foram informados dos objetivos e procedimentos do estudo e todos assinaram um termo de consentimento livre e esclarecido para participar da pesquisa. Para ambos os estudos, foi necessário realizar aquisição de imagens para as análises propostas. Nos dois casos, foi utilizada uma câmera digital (modelo: GoPro 8 Hero Black; resolução: 1920x1440 pixels; frequência de aquisição: 30 Hz) posicionada acima da quadra de basquetebol enquadrando toda a área de interesse: no caso do primeiro estudo, o circuito em que foram realizadas as tarefas e, no caso do segundo estudo, toda a quadra de basquetebol. No primeiro estudo, os vídeos foram cortados no início e no fim de cada tarefa (vide seção 4.1.2.2.). No segundo estudo, os vídeos foram cortados por quarto de jogo e foram removidos os períodos de tempo técnico e intervalo de jogo em que o cronômetro está parado e não há nenhum jogador em quadra (vide seção 4.2.2.2.). No segundo estudo, foi filmado um quarto de jogo oficial de basquetebol profissional da primeira divisão nacional, onde foram rastreados os atletas pertencentes a uma das equipes. Os procedimentos do estudo foram aprovados pelo Comitê de Ética em Pesquisa da Universidade Estadual de Londrina (parecer n. 5.626.309, CAAE: 61495422.4.0000.5231; Anexo 1).

3.2. TAREFAS ESTÁTICAS E DINÂMICAS

No primeiro estudo, testamos a acurácia estática da rede neural, onde cada um dos participantes foi instruído a permanecer parado sobre cada um dos dez pontos usados para calibração. Para cada ponto, 30 frames (1 segundo) foi analisado usando a rede neural treinada para reconhecer os pontos-chave pré-determinados. Na sequência, a posição do participante no plano da imagem foi reconstruída no sistema de referência com as coordenadas reais conhecidas. Após o teste estático, os

participantes foram submetidos às seguintes tarefas para determinar a acurácia dinâmica: caminhar, trotar, correr e a tarefa de *lane-drill*. Nas tarefas de caminhar, trotar e correr, os participantes foram instruídos a se deslocarem de frente em um percurso retangular de aproximadamente 23,8 metros delimitado por linhas no chão da quadra. Para cada tarefa, os participantes foram instruídos a aumentarem progressivamente a velocidade (atingindo a velocidade máxima na tarefa de correr). Para a tarefa de *lane-drill*, os participantes foram instruídos a se deslocarem de frente, lateralmente e de costas, o mais rápido possível, em um percurso retangular de aproximadamente 47,6 metros (ida-e-volta). Cada tarefa foi realizada duas vezes e os participantes foram instruídos a permanecerem ao máximo acima das linhas no chão que compunham o percurso das tarefas (Figura 1).

Figura 1 – Representação do percurso delimitado e das tarefas realizadas. O Painel A representa as tarefas de andar, trotar e correr, enquanto o Painel B representa a tarefa modificada de *lane-drill* (as linhas tracejadas pretas e azuis indicam a primeira e a segunda volta do percurso, respectivamente).



3.3. CALIBRAÇÃO DA ÁREA DE INTERESSE

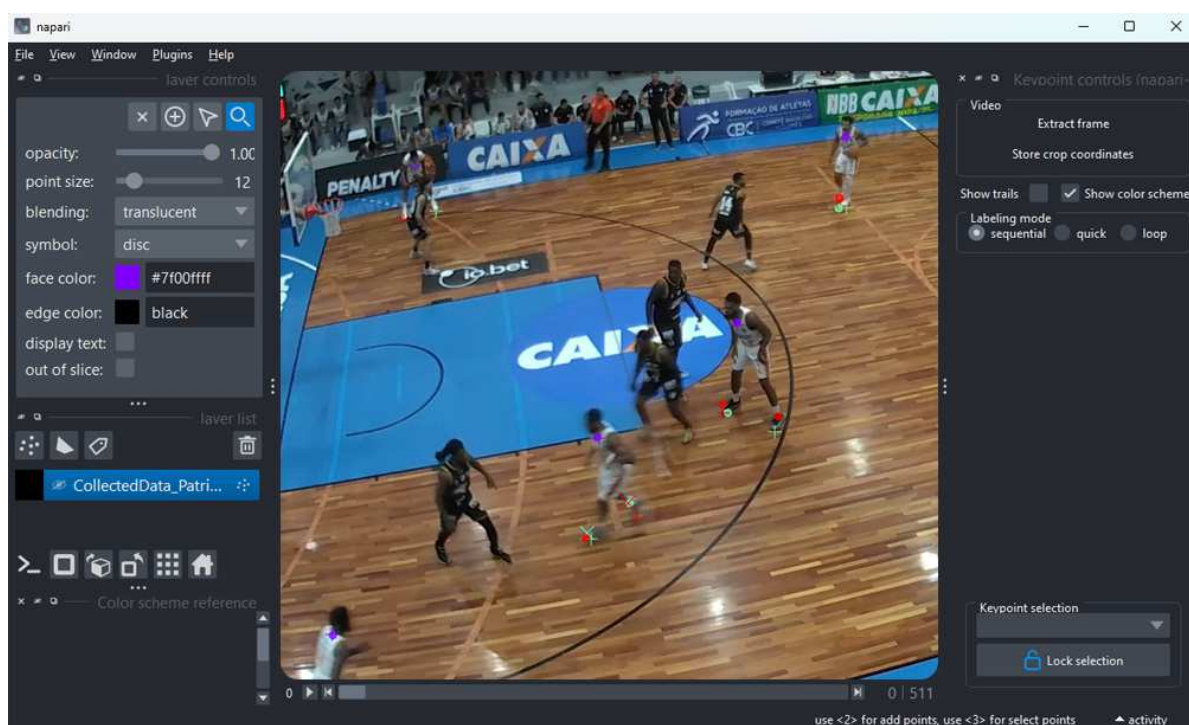
Para o primeiro estudo, a área onde foram realizadas as tarefas foi determinada a partir de dez pontos na quadra, com suas posições (coordenadas reais) e distância entre os

pontos medidas diretamente com uma trena laser (modelo: GLM 50-12, Bosch, Alemanha). Após isso, as projeções de cada ponto na imagem foram determinadas para obter as coordenadas em pixels no plano da imagem. A partir dos pontos associados ao sistema de coordenadas reais e seus correspondentes no sistema de coordenadas da imagem, os processos de homografia foram feitos pelo método de DLT - *Direct Linear Transformation* (ABDEL-AZIZ; KARARA, 2015).

3.4. TREINAMENTO DA REDE NEURAL

O treinamento da rede neural utilizada nos dois estudos foi similar, seguindo os procedimentos descritos por Lauer e colaboradores (2022), usando a versão 2.2.1 do DLC. Em ambas as situações, o primeiro passo foi extrair os frames dos vídeos que compuseram o conjunto de imagens de treinamento. Na sequência, um pesquisador rotulou manualmente três pontos-chave em cada jogador presente em cada frame do conjunto de treinamento, sendo eles: (1) esterno ou ponto médio entre as escápulas; (2) pé direito e (3) pé esquerdo. Foram utilizados 100 frames para compor o conjunto de treinamento em ambos os modelos. As etapas de extração e rotulagem dos frames foram feitas por meio da interface de usuário do DeepLabCut (Figura 2).

Figura 2 – Interface de rotulagem do conjunto de treinamento no DeepLabCut.

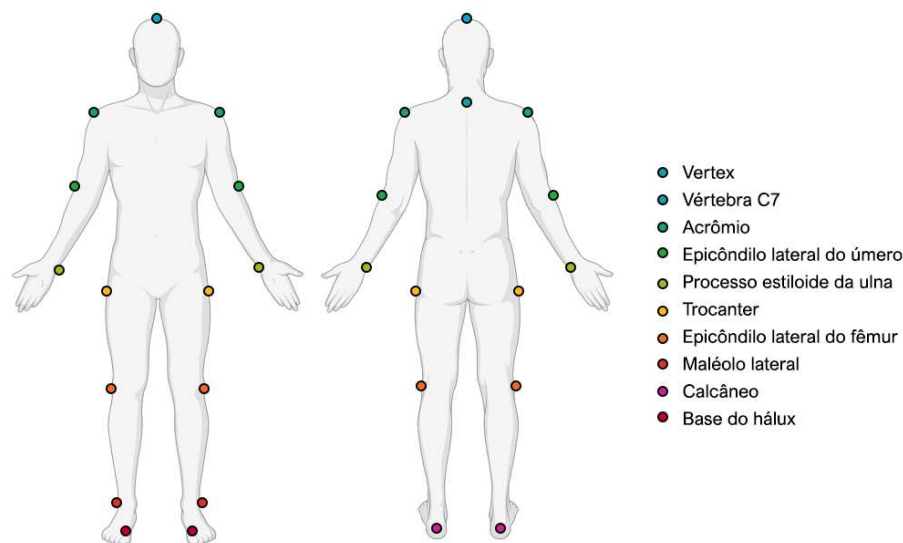


As redes neurais foram treinadas a partir dos pesos iniciais do modelo DLCRNet_ms5 para 200.000 iterações (LAUER et al., 2022), utilizando uma divisão de treino/teste de 90/10 (i.e. 90% dos frames rotulados foram utilizados para o treinamento e 10% para avaliação dos resultados do treinamento). Os parâmetros utilizados no treinamento dos modelos, como o número de pontos-chave, o número de frames rotulados e o número de iterações foram determinados a partir de testes prévios nos quais foram avaliados: a curva de aprendizagem e os erros do modelo fornecidos pelo DLC. Adotamos os parâmetros que culminaram nos menores erros de avaliação da rede neural antes de prosseguir para análise dos vídeos de interesse. No primeiro estudo, o modelo escolhido teve erros de treinamento de 1,53 pixels e erros de teste de 2,42 pixels (tamanho da imagem: 1980x1440 pixels). No segundo estudo, os erros de treinamento do modelo escolhido foram de 3,73 pixels e os erros de teste foram de 9,18 pixels (tamanho da imagem: 1980x1440 pixels). As ferramentas do DLC se baseiam em Python (v. 3.8.13) e implementa a biblioteca TensorFlow para abordagens de *deep learning*. Ambos os modelos foram treinados com uma placa de vídeo NVIDIA® GeForce RTX 3050 (NVIDIA Corp, Santa Clara, EUA).

3.5. CAPTURA OPTOELETRÔNICA

Para a captura optoeletrônica do primeiro estudo, foram utilizadas oito câmeras optoeletrônicas (OptiTrack® - Optical Motion Capture Solutions, NaturalPoint, EUA; frequência de aquisição: 120 Hz), posicionadas ao redor da área onde as tarefas ocorreram. Os participantes usaram vestimentas apropriadas para minimizar erros relacionados ao movimento dos marcadores. Dezoito marcadores reflexivos foram acoplados ao corpo dos participantes utilizando um modelo biomecânico com os seguintes pontos anatômicos: vertex, vértebra C7 e os pares bilaterais do acrômio, epicôndilo lateral do úmero, processo estiloide da ulna, trocanter, epicôndilo lateral do fêmur, maléolo lateral, calcâneo e base do hálux.

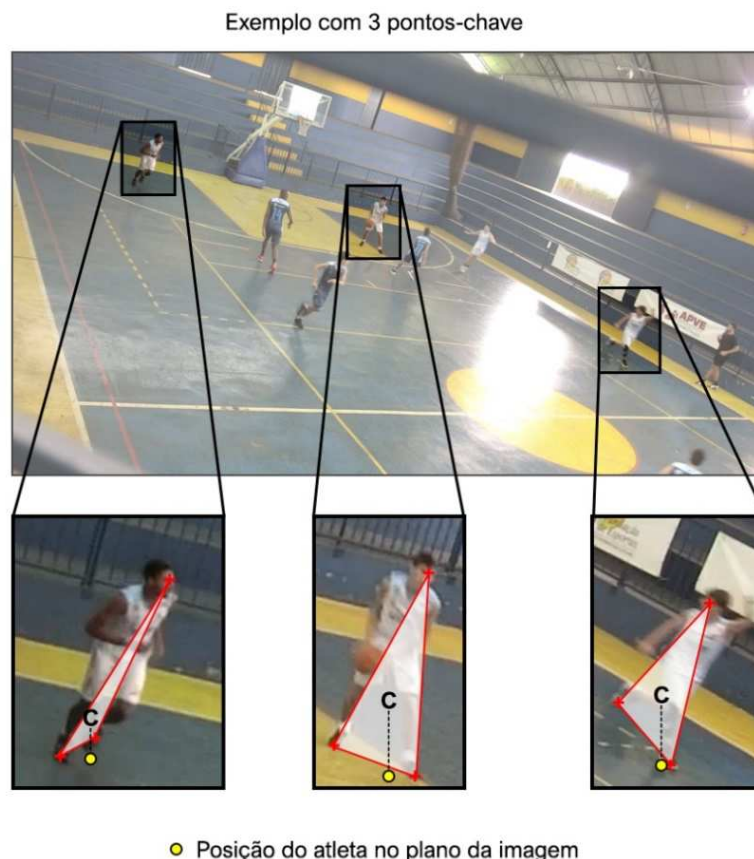
Figura 3 – Modelo biomecânico de 18 marcadores utilizado no primeiro estudo.



3.6. PROCESSAMENTO DE DADOS

No primeiro estudo, os dados obtidos a partir do DLC foram filtrados usando um filtro Butterworth passa-baixa de 3ª ordem com frequência de corte de 0,4 Hz, determinado após análise residual e espectral. No segundo estudo, foram utilizadas diferentes abordagens de processamento descritas na seção 4.2.2.3. Para determinar o ponto único que representa a posição do participante na quadra, um envoltório convexo foi desenhado utilizando as coordenadas dos pontos-chave e o centroide desta forma geométrica foi calculado em cada frame e projetado na menor coordenada Y do envoltório (vertical no plano bidimensional da imagem). Assumiu-se este ponto como a intersecção entre o participante e o chão e, portanto, foi utilizado como a medida para fins de comparação depois da reconstrução bidimensional (Figura 4).

Figura 4 – Exemplo da determinação do ponto que representa a posição do atleta no plano da imagem. C = centroide do invólucro convexo formado pelos pontos-chave rastreados pelo DLC.



No primeiro estudo, os dados obtidos a partir do OptiTrack foram filtrados usando um filtro Butterworth passa-baixa de 4ª ordem com frequência de corte de 10 Hz, determinado após análise residual e espectral. Lacunas de menores de 10 ms nos dados foram preenchidas utilizando interpolação do tipo *spline*. O centro de massa (CM) de cada participante – entendido aqui como a posição X e Y do centro de massa projetada no chão – foi utilizado como medida de referência. O CM dos participantes foi estimado usando o modelo adaptado de De Leva (1996).

No primeiro estudo, a distância percorrida total por cada participante em cada tarefa foi calculada como mostra a seguinte equação:

$$d = \sum \sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2}$$

onde d representa a distância percorrida total em cada tarefa e i representa cada ponto na série temporal. Para calcular a velocidade instantânea de cada participante em cada tarefa, a seguinte equação foi usada:

$$v_i = \frac{\sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2}}{\frac{1}{f}}$$

onde v representa a velocidade instantânea, i representa cada ponto na série temporal e f representa a frequência de aquisição de cada sistema. Posteriormente, ambos os sistemas (DLC e OptiTrack) foram sincronizados visualmente pelo início de um evento em comum (e.g. contato do calcanhar no solo) e pela duração da tarefa (indo de 0 a 100 por cento em cada tarefa), para comparação das séries temporais de velocidade.

3.7. ANÁLISE ESTATÍSTICA

No primeiro estudo, erros absolutos e relativos para cada tarefa determinaram a acurácia da distância percorrida total estimada por cada sistema. Ademais, a concordância entre sistemas foi testada por meio dos gráficos de Bland-Altman com um intervalo de confiança de 95% e por meio do coeficiente de correlação intraclass (ICC), classificado como $< 0,5$ = baixa, $0,5 - 0,74$ = moderada; $0,75 - 0,9$ = boa; e $> 0,9$ = excelente. A acurácia dos dados de velocidade foi estimada utilizando a média do erro absoluto (MEA) e os erros de root mean square (RMSE). Ainda, a técnica de mapeamento paramétrico estatístico unidimensional (SPM1D; www.spm1d.org) foi utilizada para comparar as séries temporais de velocidade entre sistemas (PATAKY et al., 2019). O nível de significância foi estabelecido em $p < 0,05$ e um mapa paramétrico estatístico (SPM {t} e o valor crítico de t) foi calculado ao longo da duração da tarefa. Em cada tarefa, as velocidades estimadas foram consideradas estatisticamente diferentes umas das outras quando o valor de SPM {t} excedeu o valor crítico de t. A SPM1D foi calculada no MATLAB (versão: 2021a, MathWorks Inc., USA) e todas as outras análises foram realizadas no software R v.4.2.2.

3.8. AVALIAÇÃO DA QUALIDADE DE RASTREAMENTO

No segundo estudo, que busca avaliar a qualidade de rastreamento do DLC em jogos oficiais, os vídeos foram analisados em função do ponto que representa a posição do jogador na quadra. Após a análise dos vídeos, foi realizado um pareamento entre as posições reais (rotuladas por usuário) e as posições estimadas pelo DLC. Na sequência, foram registrados os números de predições corretas (verdadeiros positivos; VP e verdadeiros negativos; VN) e predições incorretas (falsos positivos; FP e falsos negativos; FN). Além disso, foram registradas também as ocorrências de erros de associação de identidade dos jogadores (IDSW) – quando uma identidade é associada ao jogador incorreto. A partir desses valores, foi calculada uma série de métricas relacionadas aos erros de detecção e associação (LUITEN et al., 2021).

3.7.1. Acurácia

A acurácia é calculada como a taxa entre as amostras classificadas corretamente e o número total de amostras avaliadas. Seu valor varia de 0 a 1 – onde 1 representa a perfeita predição de todas as amostras positivas e negativas e 0 representa a predição incorreta de todas as amostras positivas e negativas. Ela pode ser calculada pela seguinte fórmula:

$$Acurácia = \frac{VP + VN}{VP + FP + VN + FN}$$

3.7.2. Recall

A medida de recall, também conhecida como sensibilidade denota a taxa de amostras positivas classificadas corretamente. É calculada como a taxa entre as amostras positivas classificadas corretamente e todas as amostras atribuídas na classe positiva. O valor de recall varia entre 0 e 1, onde 1 representa a perfeita predição de todas as amostras positivas e 0 representa a predição incorreta de todas as amostras positivas. Pode ser calculada pela seguinte fórmula:

$$Recall = \frac{VP}{VP + FN}$$

3.7.3. Especificidade

A especificidade (versão negativa do recall) denota a taxa de amostras negativas classificadas corretamente. É calculada como a taxa entre as amostras negativas classificadas corretamente e todas as amostras atribuídas na classe negativa. O valor de especificidade varia entre 0 e 1, onde 1 representa a perfeita predição de todas as amostras negativas e 0 representa a predição incorreta de todas as amostras negativas. Pode ser calculada pela seguinte fórmula:

$$\text{Especificidade} = \frac{VN}{VN + FP}$$

3.7.4. Índice de Jaccard

O índice de Jaccard é uma medida de similaridade que indica a semelhança entre dois conjuntos de dados pareados (o verdadeiro e o predito pelo algoritmo) no que diz respeito a acurácia de detecção dos objetos. É considerada também uma medida de avaliação da detecção (DetA). Pode ser calculada pela seguinte fórmula:

$$\text{Índice de Jaccard} = \frac{VP}{VP + FN + FP}$$

3.7.5. MOTA (Multiple Object Tracking Accuracy)

MOTA é uma das métricas mais comuns para avaliar a qualidade de rastreamento e ela mede três tipos de erros – os erros de detecções de FNs e FPs, bem como o erro de associação por meio dos IDSW (BERNARDIN; STIEFELHAGEN, 2008). Pode ser calculada pela seguinte fórmula:

$$\text{MOTA} = 1 - \frac{FN + FP + IDSW}{n \text{ de detecções reais}}$$

4. RESULTADOS

4.1. ESTUDO 1: RASTREAMENTO AUTOMÁTICO DE JOGADORES DE ESPORTES DE QUADRA INDOOR UTILIZANDO UMA ABORDAGEM DE DEEP LEARNING BASEADA EM VÍDEO: UM ESTUDO DE VALIDADE CONCORRENTE

Resumo

Este estudo teve como objetivo avaliar a validade concorrente do DeepLabCut (DLC) para estimar a posição e velocidade de um jogador em percurso pré-determinado em uma quadra de basquetebol. Dez jogadores jovens de basquetebol (idade: $16,5 \pm 0,5$ anos; altura: $181,7 \pm 2,4$ cm; massa corporal: $75,7 \pm 3,6$ Kg) foram submetidos a tarefas estáticas e dinâmicas. O rastreamento utilizando o DLC foi comparado a um sistema optoeletrônico. A validade foi analisada por meio de erros absolutos e relativos, análise gráfica de Bland-Altman, coeficiente de correlação intraclassa (ICC), erro quadrático médio e mapeamento paramétrico estatístico (SPM1d) para as séries temporais de velocidade. De maneira geral, os valores de ICC mostraram boa a excelente confiabilidade entre sistemas. Os erros foram maiores à medida que a velocidade proposta das tarefas aumentou. A SPM1d indicou diferenças significativas ($p < 0,05$) nas curvas de velocidade dos dois sistemas durante momentos de mudança de direção de 90 graus em velocidades mais altas. Não houve diferenças significativas entre sistemas em deslocamentos lineares ou mudanças de direção em velocidades menores. Coletivamente, nossos resultados demonstraram que o DLC pode ser considerado um método válido e custo-benefício para determinar a distância percorrida e a velocidade instantânea durante atividades em uma quadra indoor.

Palavras-chave: deep learning; cinemática; rastreamento; esportes; validação.

Este estudo está em revisão na Sports Engineering.

4.1.1. Introdução

Novas tecnologias têm impulsionado um avanço significativo na análise de desempenho esportivo ao longo dos anos, especialmente devido ao volume de informações disponíveis e sua diversidade de aplicações práticas (COLOMER et al., 2020; LORD et al., 2020). Por exemplo, o uso de tecnologias de rastreamento e análise do movimento está se tornando comum em organizações esportivas pois estas permitem análises profundas das características físicas, técnicas e táticas do treinamento e da competição (BARRIS; BUTTON, 2008; TORRES-RONDA et al., 2022). Quanto às análises da demanda física, as trajetórias dos jogadores podem ser resumidas em métricas de cinemática linear como deslocamento, velocidade e aceleração (LINKE; LINK; LAMES, 2018). Ademais, dados de rastreamento são capazes de avaliar a proficiência de passe (MERLIN et al., 2024) e deslocamentos coletivos e coordenação interpessoal podem ser utilizadas para explorar o comportamento da equipe em função de momentos-chave durante a partida (CAETANO et al., 2019; MERLIN et al., 2020; MOURA et al., 2016).

Os sistemas de rastreamento mais utilizados por organizações profissionais são sensores acoplados ao corpo dos atletas (sistemas de posicionamento global ou local, por exemplo) e em rastreamento baseado em vídeo (MCLEAN et al., 2019; TORRES-RONDA et al., 2022). Este último utiliza visão computacional e técnicas de processamento de imagens para detectar a trajetória dos atletas e tem sido aplicado em diferentes contextos esportivos (BARROS et al., 2007; DE OLIVEIRA BUENO et al., 2014; PEREIRA et al., 2017) com erros médios de medida de 0,3 a 0,7 m (FIGUEROA; LEITE; BARROS, 2006a; LARA et al., 2018; LINKE; LINK; LAMES, 2018). Entretanto, apesar dos avanços recentes nesses algoritmos, o rastreamento baseado em vídeo ainda enfrenta muitos desafios e muitos casos requerem correção manual pelo pesquisador ou profissional devido a variação nas cores de do plano de fundo, oclusões, movimentos imprevisíveis, *motion blur*, baixa resolução etc. (MANAFIFARD; EBADI; ABRISHAMI MOGHADDAM, 2017; MOURA; CAETANO; TORRES, 2025).

Abordagens que potencializam o rastreamento baseado em vídeo têm sido propostas com o avanço das redes neurais convolucionais e dos *frameworks* de *deep learning* acelerados por GPU (MOURA; CAETANO; TORRES, 2025). Estas abordagens utilizam algoritmos que buscam detectar múltiplos indivíduos em uma cena, atribuir

uma identidade a cada um e (re)associar esta identidade ao mesmo indivíduo ao longo do vídeo (LUO et al., 2021). Entre outros, alguns dois mais utilizados em estudos do movimento humano incluem o CMU-Pose (CAO, Z. et al., 2016), OpenPose (CAO, Z. et al., 2018), AlphaPose (FANG et al., 2022) e o YOLO-Pose (MAJI et al., 2022). Embora esses algoritmos não tenham sido inicialmente desenvolvidos para atender as demandas do contexto esportivo, alguns estudos exploraram seu potencial uso no badminton (CHEN; YANG; TU, 2021), padel (JAVADIHA et al., 2021), handebol (HOST; IVAŠIĆ-KOS; POBAR, 2020; HOST; POBAR; IVASIC-KOS, 2023), voleibol (TANG et al., 2024), futebol (KALAFATIC; HRKAC; BRKIC, 2022) e no basquetebol (JOHNSON, N., 2020; KARUNGARU et al., 2023; WU et al., 2019a).

Estes algoritmos utilizam redes neurais treinadas com bancos de dados de milhares de imagens rotuladas e pontos-chave padronizados. Portanto, uma alternativa que pode ser interessante para pesquisadores e profissionais da análise do movimento humano é utilizar abordagens que incorporem técnicas de *transfer learning* nos algoritmos. Dessa maneira, o usuário pode determinar quais os pontos-chave de interesse (e.g. articulações, partes do corpo, objetos etc.) e a rede neural pode “aprender” as características importantes de uma imagem sendo treinada com um pequeno conjunto de imagens nas quais o usuário rotula os locais corretos dos pontos-chave escolhidos. Assim, ao invés de características fixas ou embutidas no código, estas redes neurais flexíveis permitem ao pesquisador selecionar quais pontos-chave devem ser rastreados de acordo com diferentes objetivos de estudo.

Um desses métodos é o DeepLabCut (DLC), um pacote aberto de ferramentas que combina redes neurais profundas pré-treinadas e detecção de características customizáveis para prever pontos-chave em novos vídeos (LAUER et al., 2022; MATHIS et al., 2018). O DLC tem sido uma ferramenta emergente nas ciências do movimento, expandindo os escopos de pesquisa por permitir análises quantitativas do movimento em configurações mais ecológicas sem utilizar de equipamentos com alto custo financeiro (SEETHAPATHI et al., 2019). Estudos anteriores utilizando o DLC têm encontrado resultados promissores analisando a cinemática de diferentes tarefas como caminhar (NEEDHAM et al., 2021), correr (JOHNSON; OUTERLEYS; DAVIS, 2022; NEEDHAM et al., 2021), saltar (DRAZAN et al., 2021; NEEDHAM et al., 2021) e movimentos subaquáticos (CRONIN et al., 2019; PAPIĆ et al., 2021). Além de

incorporar técnicas de *transfer learning*, o DLC utiliza uma abordagem *bottom-up*, na qual os pontos-chave são localizados e as redes neurais treinadas são responsáveis por conectar esses pontos e determinar a pose dos indivíduos. Essa característica pode ser particularmente útil por permitir o rastreamento de atletas similares com identidades distintas em proximidade, por exemplo.

Um dos requisitos fundamentais para a utilização desses sistemas – do DLC ou de qualquer outro – para explorar a cinemática linear de jogadores no seu campo de jogo é a validação de suas estimativas comparando-as com um sistema padrão-ouro. Alguns estudos avaliaram esta validade utilizando tarefas específicas no taekwondo (BANKS et al., 2024) e no futebol (LINKE; LINK; LAMES, 2018, 2020). Esses estudos avaliaram a validade de diferentes sistemas *markerless* para estimar o deslocamento de atletas realizando tarefas específicas da modalidade em situações simuladas e tiveram resultados promissores. Entretanto, até o momento, nenhum estudo testou a validade do DLC para estimar variáveis cinemáticas no rastreamento de atletas em tarefas específicas. A obtenção acurada dessas variáveis com o DLC permitiria profissionais e pesquisadores explorarem as demandas de deslocamento durante o treinamento e competição tendo em vista um modelo de baixo custo, mais flexível, customizável e com uma interface de fácil usabilidade. Portanto, o presente estudo teve como objetivo testar a validade concorrente do DLC para estimar a posição e velocidade de jogadores durante tarefas estáticas e dinâmicas em ambiente *indoor*. É esperado que o DLC apresente validade concorrente aceitável para estimar a posição e velocidade dos jogadores em relação ao sistema padrão-ouro utilizado, com maiores erros associados a tarefas executadas em maior velocidade.

4.1.2. Método

4.1.2.1. Participantes

A amostra foi composta, por conveniência, por dez atletas jovens de basquetebol (idade: $16,5 \pm 0,5$ anos; massa: $75,7 \pm 3,6$ Kg; estatura: $181,7 \pm 2,4$ cm) pertencentes às equipes das categorias de base de Londrina – PR. Eles pertenciam à mesma equipe que participava de competições estaduais e treinavam cinco dias por semana. Antes do início da coleta de dados, os participantes e seus responsáveis legais foram informados dos objetivos e procedimentos do estudo e todos assinaram um termo de consentimento livre e esclarecido para participar da pesquisa. Os procedimentos do estudo foram aprovados pelo Comitê de Ética em Pesquisa da Universidade Estadual de Londrina (parecer n. 5.626.309, CAAE: 61495422.4.0000.5231).

4.1.2.2. Coleta de dados e calibração

Uma câmera digital (modelo: GoPro 8 Hero Black; resolução: 1920x1440 pixels; frequência de aquisição: 30 Hz) foi posicionada acima da quadra de basquetebol enquadrando a área onde foram realizadas as tarefas. A área onde foram realizadas as tarefas foi determinada a partir de dez pontos na quadra, com suas posições (coordenadas reais) e distância entre os pontos medidas diretamente com uma trena laser (modelo: GLM 50-12, Bosch, Alemanha). Após isso, as projeções de cada ponto na imagem foram determinadas para obter as coordenadas em pixels no plano da imagem. A partir dos pontos associados ao sistema de coordenadas reais e seus correspondentes no sistema de coordenadas da imagem, os processos de calibração e reconstrução bidimensional foram feitos pelo método de Direct Linear Transformation (ABDEL-AZIZ; KARARA, 2015).

4.1.2.3. Captura optoeletrônica

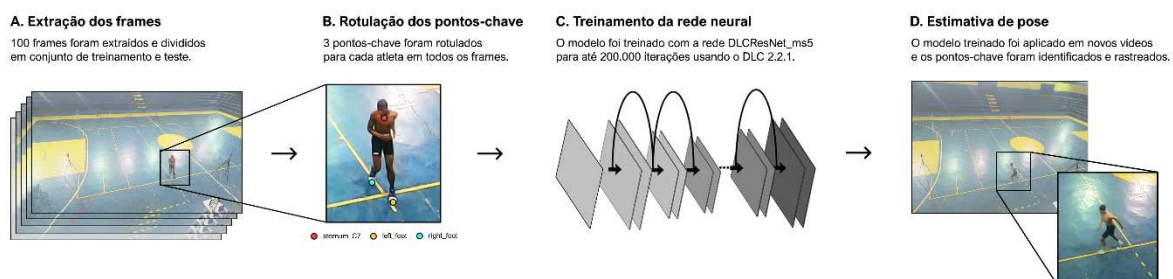
Oito câmeras optoeletrônicas (OptiTrack® - Optical Motion Capture Solutions, NaturalPoint, EUA; frequência de aquisição: 120 Hz) foram posicionadas ao redor da área onde as tarefas ocorreram. Os participantes usaram vestimentas apropriadas para minimizar erros relacionados ao movimento dos marcadores. Dezoito marcadores

reflexivos foram acoplados ao corpo dos participantes utilizando o modelo adaptado de De Leva (1996), nos seguintes pontos anatômicos: vertex, vértebra C7 e os pares bilaterais do acrômio, epicôndilo lateral do úmero, processo estiloide da ulna, trocanter, epicôndilo lateral do fêmur, maléolo lateral, calcâneo e base do hálux.

4.1.2.4. Treinamento da rede neural

O treinamento da rede neural foi realizado seguindo os procedimentos descritos por Lauer e colaboradores (2022), usando a versão 2.2.1. do DLC. O primeiro passo foi extrair os frames (dos vídeos obtidos) que irão compor o conjunto de imagens de treinamento. Na sequência, um pesquisador rotulou manualmente os três pontos-chave escolhidos em cada atleta presente em cada frame do conjunto de treinamento. Os pontos-chave escolhidos foram: (1) esterno ou ponto médio entre escápulas; (2) pé direito e (3) pé esquerdo. Vinte vídeos dos participantes realizando as tarefas foram selecionados aleatoriamente e, para cada um, 5 frames foram aleatoriamente extraídos, resultando em um total de 100 frames para compor o conjunto de treinamento. A rede neural foi treinada a partir dos pesos iniciais para detecção de características importantes do modelo DLResNet_ms5, como proposto por Lauer e colaboradores (2022). As ferramentas do DLC se baseiam em Python (v. 3.8.13) e implementa a biblioteca TensorFlow para abordagens de *deep learning*. Uma divisão percentual de 90/10 foi utilizada para o conjunto de treinamento/teste (i.e., 90% dos frames foram usados para treinamento, 10% para teste) e a rede neural foi treinada a partir de 200.000 iterações. A rede neural foi treinada usando uma placa de vídeo NVIDIA® GeForce RTX 3050 (NVIDIA Corp, Santa Clara, EUA). A Figura 1 exemplifica o processo de treinamento da rede neural utilizando o DLC.

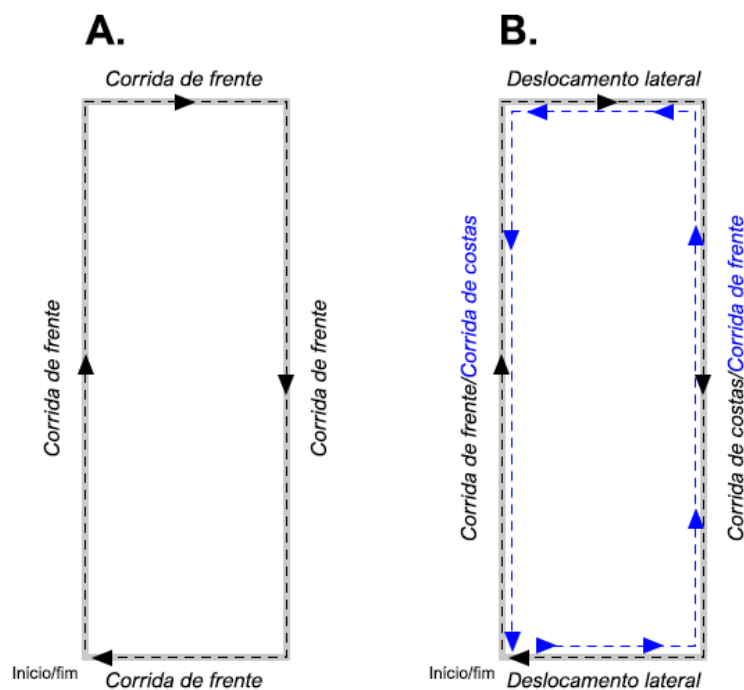
Figura 1 – Passos do treinamento da rede neural e da estimativa de pose.



4.1.2.5. Tarefas estáticas e dinâmicas

Primeiramente, testamos a acurácia estática da rede neural, onde cada um dos participantes foi instruído a permanecer parado sobre cada um dos dez pontos usados para calibração. Para cada ponto, 30 frames (1 segundo) foi analisado usando a rede neural treinada para reconhecer os pontos-chave pré-determinados. Na sequência, a posição do participante no plano da imagem foi reconstruída no sistema de referência com as coordenadas reais conhecidas. Após o teste estático, os participantes foram submetidos às seguintes tarefas para determinar a acurácia dinâmica: caminhar, trotar, correr e a tarefa de *lane-drill*. Nas tarefas de caminhar, trotar e correr, os participantes foram instruídos a se deslocarem de frente em um percurso retangular de aproximadamente 23,8 metros delimitado por linhas no chão da quadra. Para cada tarefa, os participantes foram instruídos a aumentarem progressivamente a velocidade (atingindo a velocidade máxima na tarefa de correr). Para a tarefa de *lane-drill*, os participantes foram instruídos a se deslocarem de frente, lateralmente e de costas, o mais rápido possível, em um percurso retangular de aproximadamente 47,6 metros (ida-e-volta). Cada tarefa foi realizada duas vezes e os participantes foram instruídos a permanecerem ao máximo acima das linhas no chão que compunham o percurso das tarefas.

Figura 2 – Representação do percurso delimitado e das tarefas realizadas. O Painel A representa as tarefas de andar, trotar e correr, enquanto o Painel B representa a tarefa modificada de lane-drill (as linhas tracejadas pretas e azuis indicam a primeira e a segunda volta do percurso, respectivamente).



4.1.2.6. Processamento de dados

Os dados obtidos a partir do DLC foram filtrados usando um filtro Butterworth passa-baixa de 3ª ordem com frequência de corte de 0.4 Hz, determinado após análise residual e espectral, bem como após múltiplas comparações com o sistema padrão-ouro. Para se determinar o ponto único que representa a posição do participante na quadra, um envoltório convexo foi desenhado utilizando as coordenadas dos três pontos-chave e o centroide desta forma geométrica foi calculado em cada frame e projetado na menor coordenada Y do envoltório (vertical no plano bidimensional da imagem). Assumiu-se este ponto como a intersecção entre o participante e o chão e, portanto, foi utilizado como a medida para fins de comparação depois da reconstrução bidimensional.

Os dados obtidos a partir do OptiTrack que tiveram lacunas de $1 < 10$ ms foram preenchidas utilizando interpolação do tipo *spline*. Em seguida, os dados foram filtrados usando um filtro Butterworth passa-baixa de 4ª ordem com frequência de

corde de 10 Hz, determinado após análise residual e espectral. O centro de massa (CM) de cada participante – entendido aqui como a posição XY do centro de massa projetada no chão – foi utilizado como medida de referência. O CM dos participantes foi estimado usando o modelo adaptado de De Leva (1996).

A distância percorrida total por cada participante em cada tarefa foi calculada como mostra a seguinte equação:

$$d = \sum \sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2}$$

onde d representa a distância percorrida total em cada tarefa e i representa cada ponto na série temporal. Para calcular a velocidade instantânea de cada participante em cada tarefa, a seguinte equação foi usada:

$$v_i = \frac{\sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2}}{\frac{1}{f}}$$

onde v representa a velocidade instantânea, i representa cada ponto na série temporal e f representa a frequência de aquisição de cada sistema. Posteriormente, ambos os sistemas (DLC e OptiTrack) foram sincronizados visualmente pelo início de um evento em comum (e.g. contato do calcanhar no solo) e pela duração da tarefa (indo de 0 a 100 por cento em cada tarefa), para comparação das séries temporais de velocidade.

4.1.2.7. Análise estatística

Erros absolutos e relativos para cada tarefa determinaram a acurácia da distância percorrida total estimada por cada sistema. Ademais, a concordância entre sistemas foi testada por meio dos gráficos de Bland-Altman com um intervalo de confiança de 95% e por meio do coeficiente de correlação intraclassa (ICC), classificado como $< 0,5$ = baixa, $0,5 - 0,74$ = moderada; $0,75 - 0,9$ = boa; e $> 0,9$ = excelente (KOO; LI, 2016). A acurácia dos dados de velocidade foi estimada utilizando a média do erro absoluto (MEA) e os erros de root mean square (RMSE). Para investigar se haveria diferenças em momentos específicos do circuito e complementar a medida agregada de RMSE, a técnica de mapeamento paramétrico estatístico

unidimensional (SPM1D; www.spm1d.org) foi utilizada para comparar as séries temporais de velocidade entre sistemas. O nível de significância foi estabelecido em $p < 0,05$ e um mapa paramétrico estatístico (SPM $\{t\}$ e o valor crítico de t) foi calculado ao longo da duração da tarefa. Em cada tarefa, as velocidades estimadas foram consideradas estatisticamente diferentes umas das outras quando o valor de SPM $\{t\}$ excedeu o valor crítico de t . A SPM1D foi calculada no MATLAB (versão: 2021a, MathWorks Inc., USA) e todas as outras análises foram realizadas no software R v.4.2.2.

4.1.3. Resultados

A precisão da posição dos jogadores no teste estático variou de 0,04 m (o ponto mais próximo da câmera) até 0,14 m (o ponto mais distante da câmera), como é mostrado na Figura 3.

Figura 3 – Representação o [erro de medida] da posição (em metros) de todos os dez pontos para determinar a acurácia da posição na tarefa estática. O Painel A mostra cada ponto no plano da imagem; o Painel B mostra a reconstrução bidimensional com a média e desvio padrão dos erros para cada ponto, considerando todos os participantes.



A Tabela 1 mostra os resultados gerais a respeito da distância percorrida total estimada por cada sistema. Erros médios variaram de 0,21 a 0,44 m. Os erros relativos também aumentaram à medida que a velocidade proposta da tarefa aumentou, indo de 1% (andando) à 2,1% (lane-drill). Todos os valores de ICC mostraram boa a excelente confiabilidade entre os sistemas se tratando a distância percorrida total nas

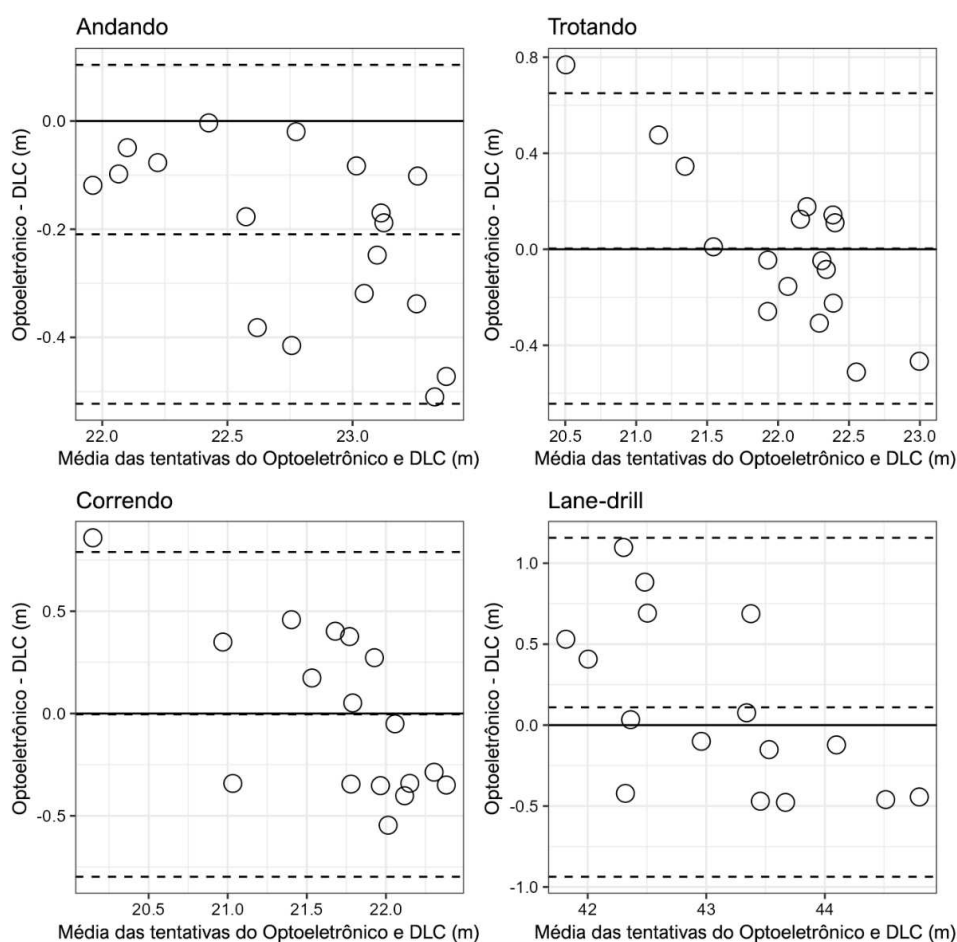
tarefas propostas. A Figura 4 mostra a concordância entre sistemas utilizando o gráfico de Bland-Altman com limites de concordância de 95%.

Tabela 1 - Concordância e confiabilidade da distância percorrida total estimada pelo DLC e pelo sistema optoeletrônico em cada tarefa.

Tarefa	Erro absoluto (m)	Erro relativo (%)	ICC [95% IC]
Andando	0,21 (0,16)	1,00 (0,76)	0,94 [0,86 – 0,98]
Trotando	0,25 (0,21)	1,20 (0,98)	0,87 [0,67 – 0,95]
Correndo	0,35 (0,18)	1,68 (0,88)	0,78 [0,50 – 0,92]
Lane-drill	0,44 (0,30)	2,11 (1,44)	0,84 [0,60 – 0,94]

Nota: erros absolutos e relativos estão expressos em média e (desvio-padrão). ICC = coeficiente correlação intraclassa (intervalos de confiança de 95% estão expressos em colchetes).

Figura 4 – Gráfico de Bland-Altman mostrando os níveis de concordância entre os sistemas para a distância percorrida total em cada tarefa. A linha tracejada do meio representa o *bias* médio estimado e as linhas tracejadas superiores e inferiores representam os limites superiores e inferiores de concordância, respectivamente.



A Tabela 2 revela os erros associados a velocidade instantânea estimada por ambos os sistemas. A MEA variou de $0,098 \pm 0,016$ (andando) até $0,332 \pm 0,047$ m·s⁻¹ (lane-drill). RMSE variou de $0,120 \pm 0,019$ (andando) até $0,415 \pm 0,071$ m·s⁻¹ (lane-drill). Semelhante à distância percorrida, os erros associados à velocidade instantânea aumentaram à medida que a velocidade proposta da tarefa aumentou.

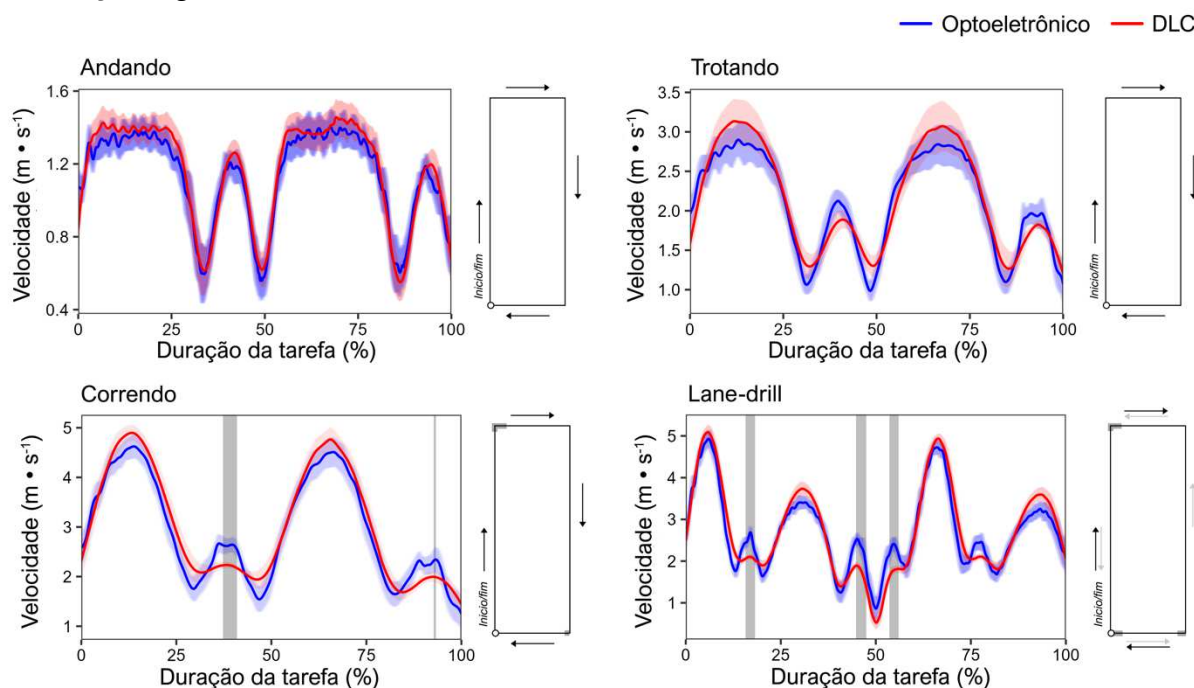
Tabela 2 – Erros associados à velocidade instantânea estimada pelo DLC e pelo sistema optoeletrônico.

Tarefa	MEA (m·s ⁻¹)	RMSE (m·s ⁻¹)
Andando	0,09 (0,02)	0,12 (0,02)
Trotando	0,21 (0,06)	0,26 (0,06)
Correndo	0,33 (0,03)	0,39 (0,04)
Lane-drill	0,33 (0,05)	0,41 (0,07)

Note: MEA (média do erro absoluto) e RMSE (*root mean square error*) estão expressos em média e (desvio-padrão).

Não houve diferença significativa entre os sistemas com relação às curvas de velocidade durante as tarefas andar ($p = 0,06$) e trotar ($p = 0,055$). Por outro lado, durante a tarefa de correr, houve dois momentos em que a curva de velocidade foi diferente entre os sistemas (isto é, onde o SPM{t} excedeu o valor crítico): entre 37-41% da duração da tarefa e entre 92-93% da duração da tarefa ($p < 0,001$). Durante a tarefa de lane-drill, as curvas de velocidade diferiram nos momentos correspondes a 16-18%, 43-47% e 53-56% da duração da tarefa ($p < 0,001$). A Figura 5 mostra a comparação entre as curvas de velocidade de cada tarefa estimadas pelo DLC e pelo sistema optoeletrônico.

Figura 5 – Série temporal da velocidade estimada pelo sistema optoeletrônico (em azul) e pelo DLC (em vermelho) nas tarefas de andar, trotar, correr e lane-drill. Os dados estão expressos em média e intervalos de confiança de 95%. As áreas sombreadas na série temporal e na representação visual do percurso indicam onde o SPM {t} excedeu o limiar crítico e, portanto, denotam os momentos em que ocorreram diferenças significativas entre os sistemas.



4.1.4. Discussão

O presente estudo teve como objetivo avaliar e comparar a acurácia do DLC com um sistema de captura de movimento considerado padrão-ouro em diferentes tarefas realizadas por jogadores de basquetebol. De maneira geral, os resultados mostraram uma validade e concordância entre os sistemas para estimar a posição do jogador, a distância percorrida total e a velocidade instantânea.

Os resultados mostraram que as medidas de acurácia para determinar a posição dos jogadores no teste estático variaram de 0,04 a 0,14 m (Figura 1). Esse resultado indica que os erros para estimar a posição aumentam à medida que a distância entre o objeto de interesse (i.e., o jogador) e a câmera aumenta e, portanto, uma calibração apropriada é necessária para atenuar esses erros. Outros estudos reportaram essa mesma dependência utilizando rastreamento baseado em vídeo, com erros variando entre 0,17 e 0,52 m (LARA et al., 2018; MOURA et al., 2012; VUČKOVIĆ

et al., 2010). Para distância percorrida total, o erro relativo médio foi de aproximadamente 1,5% - o que está alinhado com os valores de 0,2 a 1,5% reportados pela literatura utilizando outros tipos de sistemas de rastreamento baseado em vídeo (DE OLIVEIRA BUENO et al., 2014; LINKE; LINK; LAMES, 2018, 2020; MOURA et al., 2012; PEREIRA et al., 2017). Nossos resultados sugerem que os erros relativos são maiores à medida que a velocidade proposta da tarefa aumenta. Linke e colaboradores (2020) encontraram um resultado similar comparando o sistema de vídeo TRACAB's Gen 4 a um sistema de referência optoeletrônico, onde os erros relativos também foram maiores à medida que a velocidade aumentou.

Esse *trade-off* pode estar relacionado ao uso de uma única câmera, visto que nesta condição a imagem está sujeita a uma menor resolução em pixels, principalmente em jogadores distantes da câmera e, portanto, está mais suscetível a erros na determinação da posição dos jogadores (MANAFIFARD; EBADI; ABRISHAMI MOGHADDAM, 2017). Ainda, nossos resultados se mantiveram dentro de uma margem relativamente pequena e apresentaram boa a excelente confiabilidade quando comparado ao sistema de referência (Tabela 1), sugerindo que o DLC é, de uma forma geral, capaz de mensurar a distância percorrida total na quadra com acurácia suficiente e de uma maneira custo-efetiva.

Para velocidade instantânea, os resultados mostraram que o MEA médio foi de aproximadamente $0,24 \text{ m}\cdot\text{s}^{-1}$ e um RMSE médio de $0,29 \text{ m}\cdot\text{s}^{-1}$ nas diferentes tarefas. Outras investigações comparando sistemas de rastreamento baseado em vídeo com sistemas optoeletrônicos obtiveram valores de RMSE entre 0,09 e $0,79 \text{ m}\cdot\text{s}^{-1}$ (Linke et al., 2018, 2020). Além disso, a análise SPM1D mostrou diferenças significativas entre as curvas de velocidade durante momentos de mudança de direção durante as tarefas de correr e lane-drill, onde os jogadores tiveram que mudar de direção de 90 graus no percurso delimitado (Figura 2). Linke e colaboradores (2018) encontraram um resultado similar quando compararam o sistema de vídeo STATSportVU com um sistema de referência optoeletrônico, onde o RMSE da velocidade instantânea foi maior durante as mudanças de direção de 90 graus durante tarefas realizadas em alta velocidade. Outro estudo também reportou limitações avaliando o desempenho quando são mensurados tempos intermediários após mudanças consideráveis de direção ou de velocidade (LINK et al., 2019). Especificamente, nossos resultados

mostraram uma subestimação da velocidade instantânea nestes momentos, devido à suavidade da curva causada pelos métodos de filtragem aplicados aos dados posicionais brutos (LINKE; LINK; LAMES, 2020). A frequência de corte utilizada nesse estudo foi determinada após análise espectral e residual e usada como uma abordagem geral que provocou resultados aceitáveis considerando os erros posicionais (Tabela 1) e de velocidade (Tabela 2). Uma possível abordagem para solucionar esses casos poderia envolver a identificação fatores de correção e parâmetros de filtro ligados a cada sistema para minimizar erros de medida em momentos específicos (BUCHHEIT et al., 2014).

A velocidade instantânea não apresentou diferença significativa os deslocamentos lineares (de frente, laterais ou de costas) ou até mesmo nas mudanças de direção em velocidades menores (nas tarefas de andar e trotar). Portanto, nosso estudo confirma a noção de que os erros podem aumentar à medida que a velocidade do objeto de interesse aumenta. Objetos se movendo rapidamente no plano da imagem podem modificar as propriedades da imagem, tornando mais desafiador para os algoritmos de estimativa de pose detectar e rastrear os pontos-chave acuradamente (Manafifard et al., 2017). Uma possível solução seria aumentar a taxa de amostragem (e.g., para 50 ou 60 Hz), visto que os dados posicionais brutos seriam mais granulares e os cálculos das medidas derivadas (como velocidade e aceleração) poderiam ser mais precisos. Uma revisão recente mostrou que sistemas de rastreamento baseado em vídeo comumente utilizam 25-30 Hz como frequência de amostragem e que esse intervalo é suficiente para análises posicionais e espaciais (RICO-GONZÁLEZ et al., 2020a). Entretanto, eles também sugerem que diferentes variáveis de desempenho podem demandar diferentes taxas de amostragem visto que elas podem envolver diferentes velocidade (RICO-GONZÁLEZ et al., 2020a). Aumentar a frequência de amostragem requer maiores recursos computacionais e tempo de processamento. Portanto, pesquisadores e profissionais devem selecionar o cenário mais prático e eficiente de acordo com as suas limitações de equipamento e institucionais.

Nossos resultados devem ser interpretados com cuidado considerando algumas limitações do estudo: (1) a pequena área de captura limitada pelo sistema optoeletrônico. Áreas maiores permitiriam o teste de diferentes tarefas e circuitos; (2) utilizamos apenas uma câmera visando melhor custo-benefício. Melhores resultados

poderiam ser alcançados se mais câmeras fossem utilizadas com maiores resoluções de imagem (MANAFIFARD; EBADI; ABRISHAMI MOGHADDAM, 2017); (3) as tarefas selecionadas não englobam os cenários mais desafiadores para sistemas de rastreamento baseado em vídeo, como em jogos oficiais, onde os jogadores constantemente ocluem uns aos outros e trocas de identificação ocorrem. Conseqüentemente, os erros de rastreamento reportados no presente estudo não representam condições reais de competição e essa questão necessita ser explorada por futuros estudos.

4.1.5. Conclusão

Nossos resultados mostraram que o DLC pode ser considerado válido para determinar a distância percorrida e a velocidade instantânea de atletas realizando tarefas estáticas e dinâmicas em um percurso pré-delimitado em uma quadra *indoor*. Coletivamente, nossos resultados mostraram que os erros foram maiores à medida que a velocidade proposta da tarefa aumentou, especificamente durante momentos de mudança de direção de 90 graus. Embora o presente método encontre os mesmos desafios que qualquer sistema de rastreamento baseado em vídeo, o DLC oferece uma implementação flexível que pode potencialmente diminuir erros inerentes. Essa pode ser uma solução promissora para futuros pesquisadores e profissionais buscando um método de rastreamento válido e custo-benefício para esportes de quadra *indoor*.

4.2. ESTUDO 2: QUALIDADE DO RASTREAMENTO AUTOMÁTICO DE JOGADORES DE BASQUETEBOL UTILIZANDO O DEEPLABCUT™ DURANTE UM JOGO OFICIAL

Resumo

Este estudo teve como objetivo avaliar a qualidade do rastreamento automático de jogadores de basquetebol utilizando o DLC com apenas uma câmera fixada durante jogos oficiais. Para isso, foi filmado um quarto de jogo profissional de basquetebol masculino da primeira divisão nacional utilizando apenas uma câmera fixada sobre a quadra capturando toda a área de jogo. Foram testadas diferentes abordagens de processamento dos dados do DLC e avaliadas métricas de detecção e associação calculadas a partir do registro de predições corretas (verdadeiros positivos e negativos), predições incorretas (falsos positivos e negativos) e trocas de identidade. De maneira geral, a abordagem utilizando interpolação do tipo 'rloess' apresentou a melhor qualidade de rastreamento dentre as abordagens testadas – alcançando uma taxa de trocas de apenas 0,23% e um MOTA de 0,982. O tempo estimado de intervenção manual para correção das trocas em cada quarto de jogo é de aproximadamente 70 minutos. Portanto, o DLC possui uma qualidade de rastreamento automático aceitável e pode ser uma solução promissora e custo-benefício para profissionais e pesquisadores explorarem a cinemática linear de jogadores de basquetebol durante jogos oficiais.

Palavras-chave: deep learning; rastreamento de múltiplos objetos; visão computacional; basquetebol.

4.2.1. Introdução

O rastreamento automático de atletas em esportes coletivos tem recebido cada vez mais atenção devido ao seu potencial para otimizar a análise de desempenho, permitindo pesquisadores e profissionais explorarem aspectos físicos, técnicos e táticos de uma determinada modalidade (MANAFIFARD; EBADI; ABRISHAMI MOGHADDAM, 2017; MOURA; CAETANO; TORRES, 2025; TORRES-RONDA et al., 2022). Avanços na área de visão computacional e inteligência artificial tornaram possível o rastreamento por vídeo utilizando algoritmos de rastreamento de múltiplos indivíduos (CAO, W. et al., 2024; LUO et al., 2021; MOURA; CAETANO; TORRES, 2025). Seja utilizando técnicas *top-down* ou *bottom-up*, estes consistem em detectar múltiplos indivíduos em uma cena, atribuir uma identidade a cada um destes e associar esta identidade ao mesmo indivíduo ao longo do vídeo, mantendo esta associação havendo variações na posição do objeto ou nas condições da imagem (LUITEN et al., 2021; LUO et al., 2021).

Tradicionalmente, a qualidade de rastreamento destes algoritmos é avaliada utilizando grandes bases de dados não específicas do esporte, como multidões e ambientes urbanos movimentados (LEAL-TAIXÉ ET AL., 2015; RUSSAKOVSKY ET AL., 2015). Embora estes *frameworks* tenham clara aplicação no ambiente esportivo, é necessário verificar seu desempenho frente a desafios presentes nesse contexto, como a presença de oclusões frequentes, a variabilidade na iluminação e no ambiente de jogo (e.g., presença dos árbitros, torcida etc.), movimentos de alta velocidade e mudanças abruptas de direção dadas (CAO, W. ET AL., 2024; MANAFIFARD; EBADI; ABRISHAMI MOGHADDAM, 2017; RICO-GONZÁLEZ; LOS ARCOS; ET AL., 2020). Estudos anteriores exploraram o potencial uso de diferentes algoritmos no badminton (CHEN; YANG; TU, 2021), padel (JAVADIHA et al., 2021), handebol (HOST; IVAŠIĆ-KOS; POBAR, 2020; HOST; POBAR; IVASIC-KOS, 2023), voleibol (TANG et al., 2024), futebol (KALAFATIC; HRKAC; BRKIC, 2022) e basquetebol (KARUNGARU et al., 2023; WU et al., 2019a). Embora os resultados tenham sido promissores, pouca preocupação é direcionada à acurácia das medidas cinemáticas de tais algoritmos e, portanto, as interpretações decorrentes desses dados podem ser comprometidas por pesquisadores e profissionais que desejem explorar as demandas da modalidade com maior profundidade.

Um dos requisitos fundamentais para a utilização desses algoritmos por profissionais e pesquisadores é a validação de suas estimativas comparando-as com um sistema padrão-ouro. Por exemplo, o primeiro estudo desta dissertação (**vide seção 4.1**) confirmou a validade concorrente do DeepLabCut (DLC) comparando-o a um sistema optoeletrônico para estimar a distância percorrida e velocidade instantânea de atletas realizando tarefas em um circuito pré-determinado. Banks e colaboradores (2024) encontraram resultados similares avaliando a validade do algoritmo OpenPose para estimar o deslocamento de atletas de taekwondo em combates simulados. No entanto, sistemas padrão-ouro como os utilizados nos estudos citados acima apresentam limitações quanto à necessidade do uso de marcadores reflexivos, a capacidade de área capturada e a quantidade de atletas simultâneos – logo, a validação isoladamente não corresponde aos desafios do contexto esportivo.

Portanto, outro requisito fundamental para a utilização de desses algoritmos é a sua qualidade de rastreamento em contexto prático – onde os desafios mencionados acontecem (WANG, 2024). Esta é comumente avaliada por três características principais: erros de detecção, erros de associação e erros de localização (LEICHTER; KRUPKA, 2013). Erros de detecção ocorrem quando o algoritmo detecta objetos que não existem ou falha ao detectar objetos que existem na imagem. Erros de associação são aqueles em que o algoritmo atribui a identidade errada ao objeto na imagem. Erros de localização ocorrem quando a posição das detecções não está espacialmente alinhada com a posição real do objeto. Estudos anteriores que avaliam a qualidade de rastreamento de algoritmos em contexto prático no basquetebol tiveram resultados promissores (FU et al., 2020; WANG, 2024), porém a acurácia das medidas cinemáticas dos algoritmos testados ainda é desconhecida. Além disso, outros fatores que afetam a aplicação prática de profissionais e pesquisadores não foram levados em consideração como o tempo de processamento de dados, o tempo de intervenção manual e os custos logísticos e computacionais. Portanto, o objetivo deste estudo foi avaliar a qualidade do rastreamento automático de jogadores de basquetebol utilizando o DLC com apenas uma câmera fixada durante jogos oficiais.

4.2.2. Métodos

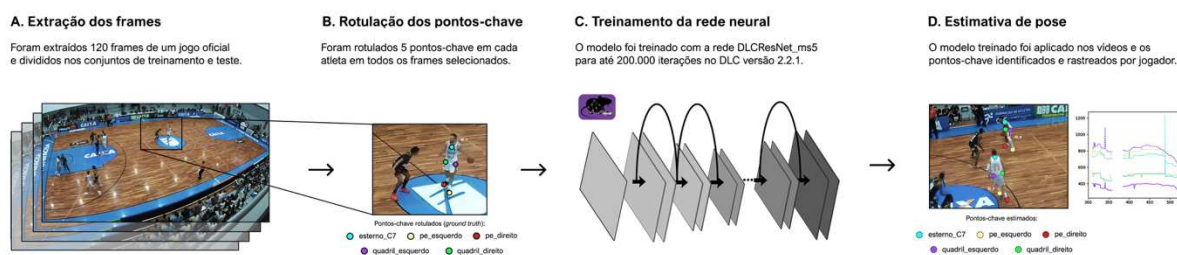
4.2.2.1. Aquisição de imagens

Uma câmera digital (modelo: GoPro 8 Hero Black; resolução: 1920x1440 pixels; frequência de aquisição: 30 Hz) foi posicionada acima da quadra de basquetebol enquadrando toda a área de jogo. Foi filmado um jogo de basquetebol profissional da primeira divisão nacional (Novo Basquete Brasil). Os vídeos foram cortados por quarto de jogo e foram removidos os períodos de tempo em que o cronômetro estava parado e não havia nenhum jogador em quadra (como em tempos técnicos e intervalo de jogo).

4.2.2.2. Treinamento da rede neural

O treinamento da rede neural foi realizado seguindo os procedimentos descritos por Lauer e colaboradores (2022), usando a versão 2.2.1. do DLC. O primeiro passo foi extrair os frames (dos vídeos obtidos) que compuseram o conjunto de imagens de treinamento. Na sequência, um pesquisador rotulou manualmente os cinco pontos-chave escolhidos em cada jogador presente em cada frame do conjunto de treinamento. Os pontos-chave foram: (1) esterno ou ponto médio entre escápulas; (2) quadril direito; (3) quadril esquerdo; (4) pé direito e (5) pé esquerdo. De todos os vídeos obtidos, 120 frames distintos foram extraídos baseados em processos de *clustering* embutidos no DLC, compondo o conjunto de treinamento. A rede neural foi treinada a partir dos pesos iniciais do modelo *resnet_50* para 200.000 iterações com uma divisão de treino/teste de 90/10 (i.e., 90% dos frames serão usados para treinamento, 10% para teste). Os erros de treinamento do modelo escolhido foram de 3,73 pixels e os erros de teste foram de 9,18 pixels (tamanho da imagem: 1980x1440 pixels). As ferramentas do DLC se baseiam em Python (v. 3.8.13) e implementa a biblioteca TensorFlow para abordagens de *deep learning*. A rede neural foi treinada usando uma placa de vídeo NVIDIA® GeForce RTX 2060 (NVIDIA Corp, Santa Clara, EUA). A Figura 2 exemplifica o processo de treinamento da rede neural utilizando o DLC.

Figura 1 – Passos do treinamento da rede neural e da estimativa dos pontos-chave.



4.2.2.3. Processamento de dados

Para determinar o ponto único que representa a posição dos jogadores na quadra, um envoltório convexo foi desenhado utilizando as coordenadas dos cinco pontos-chave e o centroide desta forma geométrica foi calculado em cada frame e projetado na menor coordenada Y do envoltório (eixo vertical no plano bidimensional da imagem). Assumimos esse ponto como a intersecção entre o jogador e o solo e, portanto, foi utilizado com medida para fins de avaliação da qualidade de rastreamento.

Realizamos a avaliação do rastreamento mediante quatro abordagens de processamento dos dados obtidos a partir do DLC. Todas as abordagens objetivaram limitar a interferência da simples troca de rótulos dos pés inviabilizasse a qualidade da identificação do atleta. Na primeira, projetamos o ponto de interesse a partir dos dados brutos de detecção do DLC sem nenhum tipo de correção ou filtro. Na segunda, aplicamos um algoritmo que visou corrigir as rotulações dos pés dos jogadores detectados – um fator importante na projeção do ponto que representa o jogador na quadra. Uma descrição do algoritmo pode ser encontrada no Quadro 1.

Quadro 1 – Descrição do algoritmo de correção da rotulação dos pés baseado em restrições de proximidade.

<p>Entrada:</p> <p>(1) Dados de cada quadro com as posições dos pés (direito e esquerdo) de todos os jogadores;</p> <p>(2) Posições dos centros dos jogadores $C_j = (x_j, y_j)$, calculados a partir dos segmentos corporais.</p>
<p>Saída:</p> <p>(1) Posições dos pés corrigidas, com rótulos inconsistentes definidos como NaN.</p>
<p>Procedimento:</p> <ol style="list-style-type: none"> 1. Para cada quadro i no conjunto de dados: 2. Calcular as posições dos centros $C_j = (x_j, y_j)$ para cada jogador $j \in \{1, 2, 3, 4, 5\}$. 3. Para cada pé $f \in \{\text{péDireito}_j, \text{péEsquerdo}_j \mid j \in \{1,2,3,4,5\}\}$: 4. Obter a posição do pé $P_f = (x_f, y_f)$. 5. Se P_f é NaN, continuar para o próximo pé. 6. Calcular as distâncias euclidianas $D = [d1, d2, d3, d4, d5]$ onde: $d_j = \text{sqrt}((x_f - x_j)^2 + (y_f - y_j)^2)$. 7. Encontrar o jogador mais próximo: $j_{\text{mais_proximo}} = \text{argmin}(D)$. 8. Determinar o jogador original j_{original} do pé f. 9. Se $j_{\text{mais_proximo}} \neq j_{\text{original}}$: 10. Contar os pés válidos (não NaN) atualmente atribuídos a $j_{\text{mais_proximo}}$. 11. Se $\text{valid_feet} \geq 2$: - Definir $P_f = (\text{NaN}, \text{NaN})$ (invalidar o pé inconsistente). 12. Caso contrário: - Manter o pé atribuído como está. 13. Fim Se. 14. Fim Para (pés). 15. Fim Para (quadros). <p>Retorno: Posições dos pés corrigidas.</p>

Na terceira abordagem, utilizamos um filtro de interpolação ‘*rloess*’ com janela amostral de 30 frames nos dados brutos de detecção do DLC e projetamos o ponto de interesse a partir dos dados filtrados. Esta técnica de suavização aplica um ajuste ponderado local, ou seja, em cada ponto dentro da janela selecionada, dando maiores pesos para os dados próximos desse ponto e menores pesos para os dados mais distantes. Dessa forma, *outliers* (pontos muito distantes do restante da janela) são ignorados sem ter influência no ajuste da curva (CUNHA, 1998). Na quarta abordagem,

unimos as duas abordagens anteriores: aplicamos o algoritmo de correção de rotulação dos pés e realizamos a interpolação com o filtro ‘*rloess*’.

4.2.2.4. Avaliação da qualidade de rastreamento

Os vídeos foram analisados em função ponto que representa a posição do jogador na quadra. Foi realizado um pareamento entre as posições reais (rotuladas por usuário) e as posições estimadas pelo DLC. Após este pareamento, foram registrados os números de predições corretas (verdadeiros positivos; VP e verdadeiros negativos; VN) e predições incorretas (falsos positivos; FP e falsos negativos; FN). Além disso, foram registradas também as ocorrências de erros de associação de identidade dos jogadores (IDSW) – quando uma identidade é associada ao jogador incorreto. Para análise da qualidade de rastreamento, foram selecionados 12.000 frames correspondentes ao início do jogo. Nesse período, houve 9 posses ofensivas e 8 posses defensivas da equipe analisada. A partir dos valores registrados nos frames analisados, foram calculadas métricas relacionadas aos erros de detecção, associação e localização:

Acurácia

A acurácia é calculada como a taxa entre as amostras classificadas corretamente e o número total de amostras avaliadas. Seu valor varia de 0 a 1 – onde 1 representa a perfeita predição de todas as amostras positivas e negativas e 0 representa a predição incorreta de todas as amostras positivas e negativas. Ela pode ser calculada pela seguinte fórmula:

$$Acurácia = \frac{VP + VN}{VP + FP + VN + FN}$$

Recall

A medida de recall, também conhecida como sensibilidade denota a taxa de amostras positivas classificadas corretamente. É calculada como a taxa entre as

amostras positivas classificadas corretamente e todas as amostras atribuídas na classe positiva. O valor de recall varia entre 0 e 1, onde 1 representa a perfeita predição de todas as amostras positivas e 0 representa a predição incorreta de todas as amostras positivas. Pode ser calculada pela seguinte fórmula:

$$\text{Recall} = \frac{VP}{VP + FN}$$

Especificidade

A especificidade (versão negativa do recall) denota a taxa de amostras negativas classificadas corretamente. É calculada como a taxa entre as amostras negativas classificadas corretamente e todas as amostras atribuídas na classe negativa. O valor de especificidade varia entre 0 e 1, onde 1 representa a perfeita predição de todas as amostras negativas e 0 representa a predição incorreta de todas as amostras negativas. Pode ser calculada pela seguinte fórmula:

$$\text{Especificidade} = \frac{VN}{VN + FP}$$

Índice de Jaccard

O índice de Jaccard é uma medida de similaridade que indica a semelhança entre dois conjuntos de dados pareados (o verdadeiro e o predito pelo algoritmo) no que diz respeito a acurácia de detecção dos objetos. Pode ser calculada pela seguinte fórmula:

$$\text{Índice de Jaccard} = \frac{VP}{VP + FN + FP}$$

MOTA (Multiple Object Tracking Accuracy)

MOTA é uma das métricas mais comuns para avaliar a qualidade de rastreamento e ela mede três tipos de erros – os erros de detecções de FNs e FPs, bem como o erro de associação por meio dos IDSW (BERNARDIN; STIEFELHAGEN, 2008). Pode ser calculada pela seguinte fórmula:

$$\text{MOTA} = 1 - \frac{FN + FP + IDSW}{n \text{ de detecções reais}}$$

4.2.3. Resultados

Com relação à frequência absoluta e relativa das predições incorretas e trocas de identidade nos frames analisados, tivemos que resultados que variaram entre 7,31 – 18,93% para a taxa de falsos positivos, 1,03 – 11,18% para a taxa de falsos negativos e 0,20 – 0,25% para a taxa de trocas de identidade (Tabela 1). Houve maiores taxas de falso positivo (18,93%) e trocas (0,25%) nos dados brutos do DLC, que não passaram por nenhum tipo de pós-processamento. A abordagem utilizando o algoritmo de correção de rotulação dos pés apresentou a maior taxa de falso negativo (11,18%).

Tabela 1 – Frequência absoluta e relativa da quantidade de predições incorretas (FP e FN) e trocas de identidade (IDSW) nos frames analisados.

Abordagem	FP (n)	FP (%)	FN (n)	FN (%)	IDSW (n)	IDSW (%)
1. Bruto	2251	18,93	781	6,57	30	0,25
2. Correção	869	7,31	1329	11,18	24	0,20
3. Interpolação	899	7,56	123	1,03	27	0,23
4. Correção + Interp.	1646	13,84	209	1,76	24	0,20

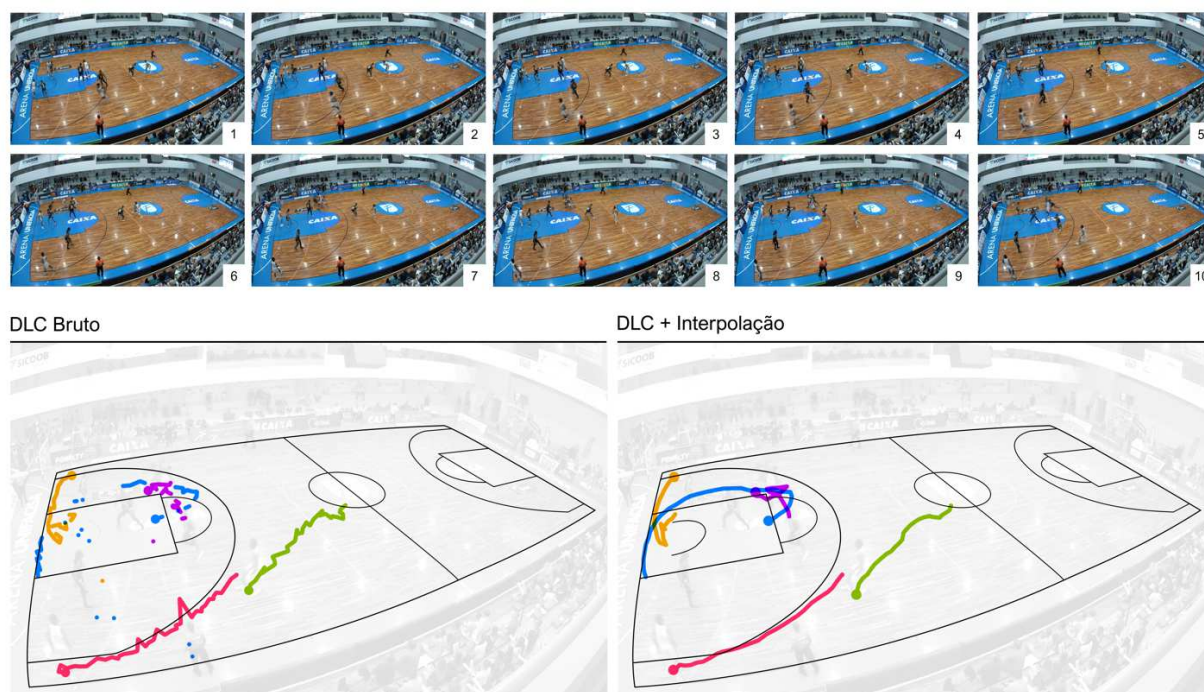
Com relação às métricas de qualidade de detecção e associação nos frames analisados, tivemos resultados que variaram entre 0,960 – 0,984 para a acurácia, entre 0,975 – 0,997 para a métrica de recall, entre 0,463 – 0,857 para a especificidade, entre 0,947 – 0,983 para o índice de Jaccard e entre 0,941 – 0,982 para a MOTA. Houve maiores valores de acurácia (0,984), de recall (0,997), do índice de Jaccard (0,983) e MOTA (0,982) na abordagem utilizando apenas a interpolação ‘rloess’. A abordagem utilizando o algoritmo de correção de rotulação dos pés apresentou o maior valor de especificidade (0,857).

Tabela 2 – Métricas da qualidade de detecção e associação nos frames analisados.

Abordagem	Acurácia	Recall	Especificidade	Índice de Jaccard	MOTA
1. Bruto	0,960	0,986	0,607	0,947	0,941
2. Correção	0,984	0,975	0,857	0,960	0,960
3. Interpolação	0,984	0,997	0,463	0,983	0,982
4. Correção + Interp.	0,972	0,996	0,466	0,969	0,967

A maioria das métricas de qualidade de detecção e associação apresentou seu maior valor na abordagem utilizando apenas a interpolação, enquanto o rastreamento utilizando apenas os dados brutos apresentou os menores valores para todas as métricas. A Figura 2 exemplifica o rastreamento da trajetória dos jogadores da equipe analisada utilizando as duas abordagens em uma sequência recortada de aproximadamente 100 frames.

Figura 2 – Demonstração das trajetórias de jogadores (equipe branca) estimadas pelas abordagens com piores resultados (dados brutos) e melhores resultados (interpolação) gerais em uma sequência recortada de 100 frames.



Foi necessária cerca de 1 hora para rotulagem do conjunto de treinamento do modelo. Além disso, nas especificações de *hardware* descritas, foram necessárias aproximadamente 8 horas para completar o treinamento do modelo e aproximadamente 1 hora para análise do vídeo de um quarto de partida. Quanto ao tempo de intervenção manual para correção das trocas de identidade, foram necessários aproximadamente 70 minutos para processar um período completo, com duração de ~17 minutos (32.275 frames). Visto que o processo de correção das trocas é feito para cada jogador, o primeiro período foi revisado quatro vezes, totalizando 129.100 frames analisados em uma taxa de processamento de 1.844 frames por

minuto. Portanto, o tempo de intervenção manual estimado para análise de uma equipe durante um jogo oficial é de aproximadamente 4 a 5 horas.

4.2.4. Discussão

O presente estudo teve como objetivo avaliar a qualidade de rastreamento automático de jogadores de basquetebol utilizando o DLC com apenas uma câmera fixada durante jogos oficiais. Os resultados indicaram uma qualidade de rastreamento aceitável para todas as abordagens de processamento utilizados, em acordo com os valores de desempenho reportados na literatura.

Os resultados mostraram uma variação de 7,31 – 18,31% na taxa de falsos positivos e de 1,03 – 11,18% na taxa de falsos negativos. Estudos anteriores que testaram diferentes algoritmos de rastreamento automático no basquetebol reportaram taxas de falsos positivos variando entre 2,9 – 33,3% e taxas de falsos negativos entre 0,1 – 97,1% (FU et al., 2020; KONG; HUANG; WANG, 2020; WU et al., 2019b; ZHANG et al., 2020). Embora nossos resultados se encontrem dentro desses limites, é justo afirmar que alguns fatores podem explicar a alta variabilidade das taxas de erro reportadas na literatura: 1) o tipo de algoritmo utilizado; 2) as condições dos vídeos analisados e a combinação entre os dois. Cada algoritmo pode apresentar características que favoreçam mais a ocorrência de falsos positivos ou de falsos negativos e a sua relação com cada tipo de erro é também dependente das condições experimentais dos vídeos utilizados nas análises.

Por exemplo, o estudo de Fu e colaboradores (2020) aplicou um modelo treinado para identificar pedestres – sem as especificidades do esporte – em vídeos de jogos recreativos (3x3 e 4x4) com uma visão limitada do espaço de jogo, com pessoas passando e outros jogos paralelos acontecendo em outras quadras ao fundo. Como esperado, a ocorrência de erros foi maior: uma taxa de falsos positivos de 33,3% e uma taxa de falsos negativos de 97,1% em ~45 mil frames analisados. Em contraponto, o estudo de Zhang e colaboradores (2020) utilizou um modelo desenvolvido especificamente para fins esportivos, denominado *DeepPlayer*, em vídeos de jogos oficiais com amplo campo de visão do espaço de jogo e teve baixas taxas de erro – sendo 1,9% para falsos positivos e 4,8% para falsos negativos em ~2,5 mil frames analisados. Portanto, o cuidado na escolha do modelo a ser utilizado e na

captura dos vídeos é de extrema importância para garantir que as taxas de erros de detecção se mantenham baixas – com menos impacto na acurácia cinemática do rastreamento.

Em relação ao número de trocas, nossos resultados mantiveram-se abaixo das taxas de 0,26 – 11,18% de todos os estudos de basquetebol encontrados até o momento (FU et al., 2020; KONG; HUANG; WANG, 2020; WU et al., 2019b; ZHANG et al., 2020). Semelhante aos erros de detecção, erros de associação também podem ser impactados pelo tipo de algoritmo e pelas condições dos vídeos analisados. Algoritmos e redes neurais podem apresentar uma robustez maior ou menor para distinguir objetos de características semelhantes, a depender do objetivo do modelo (MATHIS et al., 2020). Além disso, o número de trocas pode ser influenciado pelas condições do vídeo, especialmente em situações de oclusão dos jogadores. Alterar o campo de visão da imagem de forma que diminua o número de oclusões entre os jogadores pode facilitar o desempenho da rede neural em distinguir os jogadores. Para isso, algumas soluções são propostas, como: 1) aumentar a altura e inclinação da câmera em relação ao solo e/ou 2) aumentar o número de câmeras para capturar diferentes ângulos dos jogadores. Wu e colaboradores (2019) reportaram em seu estudo que câmeras adicionais não só diminuíram o número de trocas entre jogadores como melhorou diferentes métricas de detecção e associação. Embora câmeras adicionais representem um maior custo logístico e operacional, estas podem diminuir o tempo de intervenção manual do usuário ao lidar com menos trocas – impactando a viabilidade de aplicação do sistema de rastreamento automático (MOURA; CAETANO; TORRES, 2025). Portanto, usuários e pesquisadores devem avaliar seu contexto e decidir qual caminho é o mais viável tendo em vista o balanço entre o custo logístico e o tempo de intervenção manual necessários para analisar jogos completos.

Quanto às métricas de detecção e associação, a abordagem utilizando apenas a interpolação ‘*rloess*’ alcançou o valor mais alto na maioria das métricas avaliadas. Além disso, essa abordagem obteve um MOTA de 0,982 – maior do que os valores apresentados até então na literatura de rastreamento no basquetebol, que variaram entre 0,762 – 0,941 (FU et al., 2020; KONG; HUANG; WANG, 2020; WU et al., 2019b; ZHANG et al., 2020). A técnica ‘*rloess*’ utilizada em nosso pré-processamento dos dados funcionou com um filtro inicial para remover de maneira robusta os *outliers*

presentes nos dados brutos do nosso modelo base (causados pela elevada ocorrência de falsos positivos) e ajustar a posição dos pontos-chave em função do tempo. Por ser adaptativo e robusto, essa técnica lida muito bem com padrões complexos nos dados, como mudanças rápidas e não-lineares (CLEVELAND; DEVLIN, 1988). Estudos anteriores também utilizaram diferentes métodos de pré-processamento para lidar com as taxas de erro ou ocorrências de trocas inerentes aos modelos base utilizados. Wu e colaboradores (2019), por exemplo, testaram diferentes modelos (OpenPose, Faster-RCNN e AlphaPose) seguidos de correções padronizadas para melhorar o rastreamento e a sincronização entre múltiplas câmeras. De maneira similar, Zhang e colaboradores (2020) também testaram diferentes modelos base contra uma proposta autoral que uniu diferentes técnicas de detecção, segmentação e rastreamento dos jogadores. É importante ressaltar que cada modelo ou técnica markerless utilizada para o rastreamento automático de jogadores pode provocar erros que demandem correções específicas, como algoritmos auxiliares ou intervenções manuais para atingir os objetivos do usuário ou pesquisador.

4.2.5. Conclusão

O presente estudo demonstrou que uma abordagem de rastreamento automático utilizando o DeepLabCut como modelo base apresenta uma qualidade de rastreamento adequada e potencialmente viável considerando os custos logísticos e computacionais em jogos oficiais de basquetebol. Futuras investigações são encorajadas a explorar diferentes modelos base e algoritmos auxiliares que minimizem a necessidade de intervenções manuais, ampliando a viabilidade de aplicação de sistemas de rastreamento automático baseados em vídeo em esportes coletivos.

5. CONCLUSÃO

De maneira geral, esta dissertação pôde mostrar que o DLC: 1) possui validade concorrente adequada para estimar a distância percorrida e velocidade de jogadores realizando tarefas estáticas e dinâmicas em um percurso em uma quadra indoor e 2) apresenta uma qualidade de rastreamento aceitável quando testado a partir de uma única câmera fixada durante um jogo oficial de basquetebol. Nessas condições, o tempo estimado de intervenção manual para obter os dados ajustados do jogo todo é de aproximadamente 4 a 5 horas para um avaliador – podendo ser reduzido caso mais avaliadores (ou membros da comissão técnica) dividam a carga de trabalho.

Embora o rastreamento automático utilizando o DLC enfrente muitos dos desafios comuns dos sistemas de rastreamento baseado em vídeo, ele apresenta uma interface flexível e de fácil usabilidade, permitindo ao usuário customizar o modelo treinado de acordo com seus objetivos. Mesmo sendo testado com apenas uma câmera fixada acima da quadra, os resultados referentes à qualidade do rastreamento automático dos jogadores foram promissores e, se for viável para o usuário ou pesquisador, a adição de mais câmeras pode melhorar o desempenho do rastreamento.

Portanto, a utilização do DLC como base de um sistema de rastreamento automático de jogadores de basquetebol se mostrou altamente viável – com validade e qualidade de rastreamento adequadas, além de uma interface flexível, custos logísticos-operacionais acessíveis para a maior parte dos clubes profissionais de basquetebol e podendo entregar resultados confiáveis sobre as variáveis cinemáticas dos jogadores analisados horas após a partida. No mais, a presente dissertação contribui com a literatura nas áreas de análise do desempenho esportivo, biomecânica e visão computacional, ao testar de forma unificada tanto a acurácia cinemática das medidas obtidas pelo DLC quanto a qualidade de rastreamento do mesmo sistema aplicado em condições reais de um jogo de basquetebol.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ABDEL-AZIZ, Y. I.; KARARA, H. M. Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry. **Photogrammetric Engineering & Remote Sensing**, v. 81, n. 2, p. 103–107, 2015.

BANKS, L. et al. Accuracy of a Markerless System to Estimate the Position of Taekwondo Athletes in an Official Combat Area. **International Journal of Performance Analysis in Sport**, v. 24, n. 5, p. 479–494, 2024.

BUCHHEIT, M. et al. Integrating Different Tracking Systems in Football: Multiple Camera Semi-Automatic System, Local Position Measurement and GPS Technologies. **Journal of Sports Sciences**, v. 32, n. 20, p. 1844–1857, 2014.

CAO, W. et al. A Deep Learning Framework for Multi-object Tracking in Team Sports Videos. **IET Computer Vision**, v. 18, n. 5, p. 574–590, 2024.

CHEN, Y.T.; YANG, J.F.; TU, K.C. **Smart Badminton Detection System Based on Scaled-YOLOv4**. In: 2021 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), 2021.

CLEVELAND, W. S.; DEVLIN, S. J. Locally weighted regression: An approach to regression analysis by local fitting. **Journal of the American Statistical Association**, v. 83 n. 403, p. 596–610, 1988.

COLOMER, C. M. E. et al. Performance Analysis in Rugby Union: A Critical Systematic Review. **Sports Medicine - Open**, v. 6, n. 1, p. 4, 2020.

COLYER, S. L. et al. A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System. **Sports Medicine - Open**, v. 4, n. 1, p. 24, 2018.

CUNHA, S. A. **Metodologia para a suavização de dados biomecânicos por função não paramétrica ponderada local robusta**. Tese (doutorado) - Universidade Estadual de Campinas, Faculdade de Educação Física, Campinas, 1998.

DE LEVA, P. Adjustments to Zatsiorsky-Seluyanov's Segment Inertia Parameters. **Journal of Biomechanics**, v. 29, n. 9, p. 1223–1230, 1996.

FANG, H.S. et al. AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time. **arXiv**, 2022.

FU, X. et al. Multiple Player Tracking in Basketball Court Videos. **Journal of Real-Time Image Processing**, v. 17, n. 6, p. 1811–1828, 2020.

HOST, K.; IVAŠIĆ-KOS, M.; POBAR, M. **Tracking Handball Players with the DeepSORT Algorithm**: In: 9th International Conference on Pattern Recognition Applications and Methods, p. 593–599, 2020.

HOST, K.; POBAR, M.; IVAŠIĆ-KOS, M. Analysis of Movement and Activities of Handball Players Using Deep Neural Networks. **Journal of Imaging**, v. 9, n. 4, p. 80, 2023.

INSAFUTDINOV, E. et al. DeeperCut: **A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model**. In: Computer Vision – ECCV 2016 – Lecture Notes in Computer Science. Springer International Publishing, p. 34–50, 2016.

JAVADIHA, M. et al. Estimating Player Positions from Padel High-Angle Videos: Accuracy Comparison of Recent Computer Vision Methods. **Sensors**, v. 21, n. 10, p. 3368, 2021.

JOHNSON, C. D.; OUTERLEYS, J.; DAVIS, I. S. Agreement Between Sagittal Foot and Tibia Angles During Running Derived from an Open-Source Markerless Motion

Capture Platform and Manual Digitization. **Journal of Applied Biomechanics**, v. 38, n. 2, p. 111–116, 2022.

JOHNSON, N. **Extracting Player Tracking Data from Video Using Non-Stationary Cameras and a Combination of Computer Vision Techniques**. In: MIT Sloan Sports Analytics Conference, 2020.

KALAFATIC, Z.; HRKAC, T.; BRKIC, K. **Multiple Object Tracking for Football Game Analysis**. In: 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), p. 936–941, 2022.

KARUNGARU, S. et al. **Basketball Players Identification and Tracking using a Single Fixed Camera**. In: 17th International Conference on Signal-Image Technology and Internet-Based Systems (SITIS), p. 341–346. 2023.

KONG, L.; HUANG, D.; WANG, Y. Long-Term Action Dependence-Based Hierarchical Deep Association for Multi-Athlete Tracking in Sports Videos. **IEEE Transactions on Image Processing**, v. 29, p. 7957–7969, 2020.

LAUER, J. et al. Multi-Animal Pose Estimation, Identification and Tracking with DeepLabCut. **Nature Methods**, v. 19, n. 4, p. 496–504, 2022.

LEAL-TAIXÉ, L. et al. MOTChallenge 2015: Towards a Benchmark for Multi-Target Tracking. **arXiv**, 2015.

LEICHTER, I.; KRUPKA, E. **Monotonicity and error type differentiability in performance measures for target detection and tracking in video**. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 2003–2009. 2012.

LINKE, D.; LINK, D.; LAMES, M. Validation of Electronic Performance and Tracking Systems EPTS under Field Conditions. **PLOS ONE**, v. 13, n. 7, p. e0199519, 2018.

LORD, F. et al. Methods of Performance Analysis in Team Invasion Sports: A Systematic Review. **Journal of Sports Sciences**, v. 38, n. 20, p. 2338–2349, 2020.

LUITEN, J. et al. HOTA: A Higher Order Metric for Evaluating Multi-Object Tracking. **International Journal of Computer Vision**, v. 129, n. 2, p. 548–578, 2021.

LUO, W. et al. Multiple Object Tracking: A Literature Review. **Artificial Intelligence**, v. 293, p. 103448, 2021.

MAJI, D. et al. **YOLO-Pose: Enhancing YOLO for Multi Person Pose Estimation Using Object Keypoint Similarity Loss**. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), p. 2636–2645, 2022.

MANAFIFARD, M.; EBADI, H.; ABRISHAMI MOGHADDAM, H. A Survey on Player Tracking in Soccer Videos. *Computer Vision and Image Understanding*, v. 159, p. 19–46, jun. 2017.

MATHIS, A. et al. A Primer on Motion Capture with Deep Learning: Principles, Pitfalls, and Perspectives. **Neuron**, v. 108, n. 1, p. 44–65, 2020.

MATHIS, A. et al. DeepLabCut: Markerless Pose Estimation of User-Defined Body Parts with Deep Learning. **Nature Neuroscience**, v. 21, n. 9, p. 1281–1289, 2018.

MONTEIRO, R. L. M. et al. Enhancing Soccer Goalkeepers Penalty Dive Kinematics with Instructional Video and Laterality Insights in Field Conditions. **Scientific Reports**, v. 14, n. 1, p. 10225, 2024.

MOURA, F. A.; CAETANO, F. G.; TORRES, R. Tracking Methods in Sport: Review of Advances, Quality, and Challenges in Performance Data. **International Journal of Sports Medicine**, 2025.

NEEDHAM, L. et al. The Accuracy of Several Pose Estimation Methods for 3D Joint Centre Localization. **Scientific Reports**, v. 11, n. 1, p. 20673, 2021.

PALUCCI VIEIRA, L. H. et al. Automatic Markerless Motion Detector Method against Traditional Digitization for 3-Dimensional Movement Kinematic Analysis of Ball Kicking in Soccer Field Context. **International Journal of Environmental Research and Public Health**, v. 19, n. 3, p. 1179, 2022.

PANG, H.; ZHANG, D. Actions of Basketball Players Integrating Computer Vision and Sports Biomechanics. **Journal of Computational Methods in Sciences and Engineering**, p. 1–15, 2025.

PATAKY, T. C. et al. On the Validity of Statistical Parametric Mapping for Nonuniformly and Heterogeneously Smooth One-Dimensional Biomechanical Data. **Journal of Biomechanics**, v. 91, p. 114–123, 2019.

RICO-GONZÁLEZ, M.; PINO-ORTEGA, J.; et al. Past, Present, and Future of the Technological Tracking Methods to Assess Tactical Variables in Team Sports: A Systematic Review. **Journal of Sports Engineering and Technology**, v. 234, n. 4, p. 281–290, 2020.

RICO-GONZÁLEZ, M.; LOS ARCOS, A.; et al. The Use of Technology and Sampling Frequency to Measure Variables of Tactical Positioning in Team Sports: A Systematic Review. **Research in Sports Medicine**, v. 28, n. 2, p. 279–292, 2020.

RUSSAKOVSKY, O. et al. ImageNet Large Scale Visual Recognition Challenge. **International Journal of Computer Vision**, v. 115, n. 3, p. 211–252, 2015.

SEETHAPATHI, N. et al. Movement science needs different pose tracking algorithms. **arXiv**, 2019.

TAN, J. Q. J.; CHOW, J. Y.; KOMAR, J. The Utility of Markerless Motion Capture for Performance Analysis in Racket Sports. **Journal of Sports Engineering and Technology**, p. 17543371241230731, 2024.

TANG, G. et al. **Tracking Players in Volleyball Matches Using Vol-Bot-SORT**. In: 5th International Conference on Computer Engineering and Application (ICCEA), p. 1247–1250, 2024.

TORRES-RONDA, L. et al. Tracking Systems in Team Sports: A Narrative Review of Applications of the Data and Sport Specific Analysis. **Sports Medicine - Open**, v. 8, n. 1, p. 15, 2022.

WANG, T. Development of a Multi-Level Feature Fusion Model for Basketball Player Trajectory Tracking. **Systems and Software Computing**, v. 6, p. 200119, 2024.

WANI, M.; BHAT, F. A.; AFZAL, S.; KHAN, A. I. **Advances in Deep Learning**. 1^a ed. Springer Singapore, 2019.

WU, K.H. et al. **Robust Basketball Player Tracking Based on a Hybrid Detection Grouping Framework for Overlapping Cameras**. In: IEEE International Conference on Big Data, p. 5094–5100. 2019.f

ZHANG, R. et al. Multi-Camera Multi-Player Tracking with Deep Player Identification in Sports Video. **Pattern Recognition**, v. 102, p. 107260, 2020.

ANEXO 1



Comitê de Ética em
Pesquisa Envolvendo
Seres Humanos

UNIVERSIDADE ESTADUAL DE
LONDRINA - UEL



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: AVALIAÇÃO DA ACURÁCIA DE UM SISTEMA BASEADO EM INTELIGÊNCIA ARTIFICIAL PARA RASTREAMENTO AUTOMÁTICO DE ATLETAS

Pesquisador: BRUNO GIOVANINI

Área Temática:

Versão: 2

CAAE: 61495422.4.0000.5231

Instituição Proponente: CEFE - PROGRAMA DE PÓS - GRADUAÇÃO EM EDUCAÇÃO FÍSICA UEM/UEL

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 5.626.309

Apresentação do Projeto:

Segundo autor, novas tecnologias estão constantemente emergindo e permitindo um crescimento exponencial na análise de desempenho esportivo ao longo dos anos, como por exemplo as tecnologias de rastreamento: comuns em organizações de esportes coletivos pois estas permitem análises profundas das características físicas e táticas do treinamento e da competição, como deslocamento, velocidade e aceleração dos atletas.

Os sistemas de rastreamento mais utilizados por organizações profissionais são sensores acoplados ao corpo dos atletas (sistemas de posicionamento global ou local, por exemplo) e em rastreamento baseado em vídeo. As limitações desse método podem ser superadas utilizando uma abordagem baseada em algoritmos de estimativa de pose com deep learning. Um desses métodos é o DeepLabCut (DLC), um pacote aberto de ferramentas que combina redes neurais pré-treinadas e detecção de características customizáveis para predizer pontos-chave em novos vídeos.

Objetivo da Pesquisa:

Segundo autor, o objetivo é avaliar a acurácia do sistema DeepLabCut para rastreamento automático da trajetória de atletas de diferentes modalidades esportivas.

Endereço: LABESC - Sala 14

Bairro: Campus Universitário

UF: PR

Município: LONDRINA

CEP: 86.057-970

Telefone: (43)3371-5455

E-mail: cep268@uel.br



Comitê de Ética em
Pesquisa Envolvendo
Seres Humanos

UNIVERSIDADE ESTADUAL DE
LONDRINA - UEL



Continuação do Parecer: 5.626.309

Metodologia

Participarão deste estudo 100 atletas recreacionais, com idade entre 18 e 35 anos de idade, de sete modalidades esportivas: futebol, futsal, handebol, basquetebol, tênis e badminton. Para cada modalidade, serão realizados os mesmos procedimentos: coleta de dados e calibração, treinamento da rede neural, testes de acurácia de rastreamento (estático e dinâmico), processamento de dados, e aplicação do método em jogos oficiais ou extraoficiais.

Avaliação dos Riscos e Benefícios:

Segundo autor do projeto, os riscos são mínimos. Porém, como qualquer prática de exercício físico a participante pode sofrer uma lesão durante a execução. Portanto, caso seja necessário atendimento médico, o (a) participante será acompanhado (a) pelo pesquisador principal.

O benefício esperado será ampliar a possibilidade de realizar investigações fidedignas em diferentes competições com diferentes atletas com um sistema financeiramente acessível.

Comentários e Considerações sobre a Pesquisa:

Não há

Considerações sobre os Termos de apresentação obrigatória:

- Folha de Rosto devidamente assinada.
- Termo de confidencialidade apresentado.
- TCLE apresentado na forma de convite, com todos os elementos obrigatórios conforme Resolução CNS 466/2012.
- Cronograma corretamente apresentado, com coleta entre 01/07/2023 a 31/12/2023.
- Orçamento apresentado, no valor de R\$ 11.744,00.

Recomendações:

Em atendimento à Resolução CNS 466/2012, o suporte médico oferecido aos participantes (em caso de necessidade) não deve onerar os sistemas públicos de saúde.

Endereço: LABESC - Sala 14

Bairro: Campus Universitário

UF: PR

Município: LONDRINA

CEP: 86.057-970

Telefone: (43)3371-5455

E-mail: cep268@uel.br



Comitê de Ética em
Pesquisa Envolvendo
Serres Humanos

UNIVERSIDADE ESTADUAL DE
LONDRINA - UEL



Continuação do Parecer: 5.626.309

Conclusões ou Pendências e Lista de Inadequações:

Pendência 1: Esclarecer no item Metodologia, a forma de contato com os possíveis participantes da pesquisa, ou seja, se convite para participação na pesquisa será realizado através de contato telefônico, email ou aplicativo de mensagens (nesse caso é necessário indicar a origem dessas informações de contato e incluir termo de sigilo sobre as mesmas), se os participantes pertencem a alguma instituição/agremiação (ex. clube ou academia) é necessário incluir o termo de coparticipação para a referida instituição/agremiação.

Pendência totalmente atendida:

Forma de contato, direto com o participante.

Termo de Coparticipação apresentado.

Pendência 2: Explicitar no TCLE a informação de que serão realizadas gravações em vídeo dos participantes, para que os mesmos tenham ciência dessa ação.

Pendência totalmente atendida:

TCLE apresenta a informação de gravação em vídeo.

Pendência 3: Incluir na item Metodologia, o formato de apresentação dos resultados da pesquisa, por exemplo, dados em forma de planilha, imagens dos atletas, etc. Essa resposta pode implicar em outros desdobramentos, como por exemplo, se utilizar imagens dos participantes, deve-se indicar qua imagem serão alterada de modo a não permitir a identificação do participante.

Pendência totalmente atendida:

Forma de apresentação de resultados, através de resumo de informações, sem a inclusão de imagens dos participantes.

Pendência 4: Informar o local onde serão realizadas os teste estáticos e dinâmicos (essa informação deve constar na Metodologia e no TCLE).

Pendência totalmente atendida:

Local apresentado pelo pesquisador: Centro de Educação Física e Esportes da Universidade Estadual de Londrina.

Considerações Finais a critério do CEP:

Prezado(a) Pesquisador(a),

Endereço: LABESC - Sala 14

Bairro: Campus Universitário

UF: PR

Município: LONDRINA

CEP: 86.057-970

Telefone: (43)3371-5455

E-mail: cep268@uel.br



Comitê de Ética em
Pesquisas Envolvendo
Seres Humanos

UNIVERSIDADE ESTADUAL DE
LONDRINA - UEL



Continuação do Parecer: 5.626.309

Este é seu parecer final de aprovação, vinculado ao Comitê de Ética em Pesquisas Envolvendo Seres Humanos da Universidade Estadual de Londrina. É sua responsabilidade apresentá-lo aos órgãos e/ou instituições pertinentes.

Ressaltamos, para início da pesquisa, as seguintes atribuições do pesquisador, conforme Resolução CNS 466/2012 e 510/2016:

A responsabilidade do pesquisador é indelegável e indeclinável e compreende os aspectos éticos e legais, cabendo-lhe:

- conduzir o processo de Consentimento e de Assentimento Livre e Esclarecido;
- apresentar dados solicitados pelo sistema CEP/CONEP a qualquer momento;
- desenvolver o projeto conforme delineado, justificando, quando ocorridas, a sua mudança ou interrupção;
- elaborar e apresentar os relatórios parciais e final;
- manter os dados da pesquisa em arquivo, físico ou digital, sob sua guarda e responsabilidade, por um período mínimo de 5 (cinco) anos após o término da pesquisa;
- encaminhar os resultados da pesquisa para publicação, com os devidos créditos aos pesquisadores e pessoal técnico integrante do projeto;
- justificar fundamentadamente, perante o sistema CEP/CONEP, interrupção do projeto ou a não publicação dos resultados.

Coordenação CEP/UEL.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1997569.pdf	01/09/2022 12:20:34		Aceito
Outros	TERMO_CONFIDENCIALIDADE_BG.pdf	01/09/2022 12:20:08	BRUNO GIOVANINI	Aceito
Declaração de concordância	DECLARACAO_COPARTICIPANTE_BG_Julio_APVE.pdf	01/09/2022 12:19:35	BRUNO GIOVANINI	Aceito
Outros	RESPOSTA_PARECER_BG.docx	01/09/2022 12:19:01	BRUNO GIOVANINI	Aceito
Projeto Detalhado / Brochura Investigador	BrunoGiovanini_Projeto.docx	01/09/2022 12:14:46	BRUNO GIOVANINI	Aceito

Endereço: LABESC - Sala 14

Bairro: Campus Universitário

CEP: 86.057-970

UF: PR

Município: LONDRINA

Telefone: (43)3371-5455

E-mail: cep268@uel.br



Conselho de Ética em
Pesquisa Envolvendo
Serres Humanos

UNIVERSIDADE ESTADUAL DE
LONDRINA - UEL



Continuação do Parecer: 5.626.309

TCLE / Termos de Assentimento / Justificativa de Ausência	BrunoGiovanini_TCLE.docx	01/09/2022 12:14:32	BRUNO GIOVANINI	Aceito
Folha de Rosto	BrunoGiovanini_Projeto_FolhadeRosto.pdf	11/08/2022 15:26:17	BRUNO GIOVANINI	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

LONDRINA, 05 de Setembro de 2022

Assinado por:
Adriana Lourenço Soares Russo
(Coordenador(a))

Endereço: LABESC - Sala 14

Bairro: Campus Universitário

UF: PR

Município: LONDRINA

CEP: 86.057-970

Telefone: (43)3371-5455

E-mail: cep268@uel.br