|  |  |
| --- | --- |
| CURSO DE CIÊNCIA DA COMPUTAÇÃO – TCC | |
| (  ) PRÉ-PROJETO     ( X ) PROJETO | ANO/SEMESTRE: 2024/1 |

Previsão da Direção de Chutes em Pênaltis utilizando Estimação de Pose Humana e aprendizado de máquina

José Henrique Lenz Pellet

Prof. Aurélio Faustino Hoppe – Orientador

# Introdução

Ao longo do tempo, o futebol tem se transformado em um esporte cada vez mais competitivo, no qual cada detalhe pode significar a diferença entre a vitória e a derrota. Diante disso, de acordo com Curley e Roeder (2016), a análise dos comportamentos e jogadas dos jogadores tornou-se uma prática fundamental para equipes e atletas em busca do alto desempenho.

Curley e Roeder (2016) também afirmam que a disponibilidade de ferramentas avançadas de análise de dados, como a visão computacional e o aprendizado de máquina, tem revolucionado a forma como essas análises são realizadas, permitindo uma compreensão mais profunda dos padrões de jogo e possibilitando o desenvolvimento de estratégias mais eficazes. Por exemplo, Cui *et al*. (2018) exploraram o potencial das redes neurais convolucionais na identificação e classificação de padrões de movimento em vídeos de futebol, enquanto Memar *et al*. (2020) desenvolveram um sistema que utiliza visão computacional e aprendizado profundo para prever a direção dos chutes em cobranças de pênaltis.

Além das pesquisas acadêmicas, Curley e Roeder (2016), também apontam que existem várias ferramentas comerciais para análise de dados de movimento humano no esporte. Segundo os autores, uma dessas ferramentas é o sistema de análise de vídeo Dartfish, amplamente utilizado por equipes esportivas profissionais. O Dartfish permite aos usuários capturar vídeos de eventos esportivos, analisar o desempenho dos atletas por meio de ferramentas de marcação e anotação, e extrair métricas importantes relacionadas ao movimento e técnica (DARTFISH, 1999). Outra ferramenta comercial popular é o Prozone Sports, que oferece uma plataforma abrangente para análise de desempenho esportivo (PROZONE, 1995). O Prozone Sports combina análise de vídeo com tecnologias avançadas de rastreamento de movimento para fornecer informações sobre o desempenho dos jogadores. Segundo Curley e Roeder (2016), essas ferramentas oferecem uma solução completa para equipes e treinadores interessados em aprimorar o desempenho de seus atletas por meio da análise de dados de movimento. Além disso, empresas como a Catapult Sports e a STATSports também oferecem sistemas de monitoramento *wearable* que permitem a coleta de dados biométricos e de movimento durante treinamentos e jogos.

Segundo Bridgeman *et al*. (2019), tais ferramentas não apenas oferecem uma vantagem estratégica para equipes e atletas, mas também contribuem para o avanço do conhecimento científico sobre o movimento humano e seu desempenho no contexto esportivo. À medida que essas ferramentas e técnicas se tornam mais acessíveis, espera-se que a análise de dados de movimento continue a se expandir para novas áreas do esporte. Além das situações de jogo, os dribles, passes e marcação, estão se tornando alvos de estudo e análise. Essa abordagem holística, segundo os autores, não só aprimora o desempenho individual dos jogadores, mas também permite uma compreensão mais profunda das dinâmicas coletivas e táticas dentro de uma partida de futebol.

Pinheiro (2022) destacam que a cobrança de pênaltis, um dos momentos mais tensos e decisivos do jogo, não escapou dessa análise. A busca por compreender e prever o movimento do jogador e da bola durante uma cobrança de pênalti tornou-se um objetivo ambicioso, mas crucial, para equipes e pesquisadores interessados em maximizar suas chances de sucesso. Na análise da cobrança de pênaltis, segundo os autores, a identificação da posição corporal do jogador é um aspecto fundamental. Para Kleina *et al*. (2021), através do uso de técnicas avançadas de visão computacional e aprendizado de máquina, é possível analisar meticulosamente a postura do jogador, desde a posição dos pés até a orientação do tronco e dos membros superiores. Tal análise, conforme os autores, também permite prever a direção do chute, assim como avaliar a confiança e a técnica do jogador no momento da cobrança.

De acordo com Zhang *et al*. (2021), nesse processo de análise, as redes neurais convolucionais têm se destacado, sendo capazes de reconhecer padrões sutis na postura do jogador que podem influenciar o resultado da cobrança. Ao treinar esses modelos com um grande volume de dados de cobranças de pênaltis, segundo os autores, é possível desenvolver algoritmos precisos que podem identificar os sinais indicativos de uma cobrança bem-sucedida ou de uma possível falha. Ainda segundo Zhang *et al*. (2021), essa análise aprofundada não só oferece informações importantes para os treinadores e analistas durante os jogos, mas também pode ser utilizada para o aprimoramento individual dos jogadores, ajudando-os a aperfeiçoar sua técnica e aumentar suas chances de sucesso em momentos decisivos.

Diante deste contexto, este trabalho apresenta a seguinte pergunta de pesquisa: “De que forma a combinação da estimação de pose humana com técnicas de aprendizado de máquina pode ser empregada para prever, em tempo real, a direção de cobranças em pênaltis”.

## OBJETIVOS

O objetivo principal deste trabalho é criar um artefato computacional que detecte a pose do jogador na cobrança de pênalti, de modo a identificar o momento do contato entre o pé e a bola, bem como prever a direção de pênaltis em tempo real.

Os objetivos específicos são:

1. empregar técnicas avançadas de aprendizado de máquina e visão computacional para prever a pose, movimentos do jogador e a trajetória da bola durante as cobranças de pênalti;
2. disponibilizar um mecanismo para visualização dos dados e das previsões de direção dos chutes;
3. quantificar aspectos específicos do desempenho dos jogadores durante as cobranças de pênaltis, tais como a precisão dos chutes e a eficácia das defesas, utilizando os dados de rastreamento e análise da pose do jogador como base;
4. realizar uma análise qualitativa e quantitativa dos resultados obtidos pelo modelo, identificando suas principais vantagens e limitações em relação a abordagens tradicionais de análise de desempenho.

# trabalhos correlatos

Nesta seção serão apresentados três trabalhos correlatos com semelhanças à proposta deste projeto. A seção 2.1 discute o estudo de Pinheiro (2022), que combina a estimação de pose humana com análise notacional para avaliar cobranças de pênalti. A seção 2.2 descreve o trabalho de Silva (2022), que utiliza visão computacional para detectar o momento exato de contato do pé com a bola durante pênaltis, visando prever a direção do chute. Por fim, a seção 2.3 discute a pesquisa de Bridgeman *et al*. (2019), que se concentra na estimação e rastreamento de pose 3D em ambientes esportivos, superando desafios como movimentos rápidos e interações físicas entre atletas.

## ESTIMAÇÃO DE POSE CORPORAL INTEGRADA À ANÁLISE NOTACIONAL: UMA NOVA ABORDAGEM PARA ANALISAR ESTRATÉGIAS DE CHUTES DE PÊNALTI NO FUTEBOL DE ELITE.

Pinheiro (2022) explorou a intersecção entre a tecnologia de estimação de pose humana, com foco no uso do OpenPose e a análise notacional para avaliar as estratégias adotadas em momentos decisivos de cobranças de pênalti em competições de futebol de alto nível. Com o objetivo de detalhar os ângulos de orientação corporal dos jogadores e associá-los às táticas de ataque e defesa, incorporando a análise de 34 vídeos de cobranças de pênalti, coletados de transmissões televisivas das principais ligas europeias entre os anos de 2017 e 2020.

Para a análise dos vídeos, Pinheiro (2022) utilizou a técnica OpenPose, uma solução avançada de visão computacional, que permite a extração precisa de informações sobre a orientação dos ombros, quadris e pés dos jogadores no momento antes do chute. Segundo o autor, esta abordagem forneceu uma compreensão detalhada das dinâmicas envolvidas nas cobranças de pênalti, evidenciando como a postura corporal dos jogadores influencia diretamente nas estratégias de chute adotadas. A Figura 1 exemplifica a análise corporal dos jogadores envolvidos na batida de pênalti.

Figura 1 – Análise batedor e goleiro em pênaltis, focando em orientação corporal e movimento dos pés

|  |  |
| --- | --- |
| Uma imagem contendo Interface gráfica do usuário  Descrição gerada automaticamente | Uma imagem contendo Interface gráfica do usuário  Descrição gerada automaticamente |

Fonte: Pinheiro (2022).

Pinheiro (2022) também calculou a velocidade da bola utilizando o software Kinovea, complementando os dados de pose com métricas dinâmicas do chute. Esse método, segundo o autor, permitiu identificação de padrões nas estratégias de cobrança de pênalti, destacando a eficácia do OpenPose na análise de desempenho.

Na fase de resultados, Pinheiro (2022) observou que o ângulo dos ombros e quadris dos jogadores no momento que precede o chute, forneceram indicativos claros da direção da bola. Por exemplo, uma inclinação de 17° nos ombros em relação à direção do gol mostra um indício da escolha da direção do chute, com uma precisão de 82% nos casos analisados. No entanto, ao realizar a análise notacional das respostas dos goleiros frente aos chutes, Pinheiro (2022) observou que os goleiros que anteciparam o movimento em 0,7 segundos antes do chute, conseguiram um índice de sucesso na defesa de 70%, em comparação aos que reagiram mais tarde, cujo índice caiu para 45%.

Ainda de acordo com Pinheiro (2022), tais descobertas resultam em implicações práticas para treinadores e analistas esportivos, sugerindo que o treinamento focado na interpretação da linguagem corporal pode melhorar a capacidade dos goleiros em antecipar os chutes. Por outro lado, os batedores podem beneficiar-se ao compreender como atenuar sinais visuais que denunciem suas intenções de chute.

Por fim, Pinheiro (2022) afirma que o estudo amplia o entendimento atual sobre análise de desempenho no futebol, ao integrar tecnologias que estimam a pose com análise notacional. O autor também sugere a aplicação dessas tecnologias em outras situações de jogo, com o objetivo de compreender de forma mais ampla as dinâmicas de jogo e fornece novos recursos para o treinamento e preparação de atletas em competições de alto nível.

## Explorando a Captação de Movimento dos Jogadores de Futebol em Cobranças de Pênaltis

Silva (2022) desenvolveu um sistema que utiliza algoritmos de visão computacional para identificar o momento exato do contato do pé do jogador com a bola durante a cobrança de pênalti. Inicialmente, o estudo enfatizou a necessidade de um pré-processamento detalhado dos dados, incorporando a segmentação e captura de vídeos e imagens de pênaltis sob variadas condições de iluminação e de diferentes ângulos. Segundo o autor, este processo foi essencial para desenvolver um banco de dados sólido, propício para análises aprofundadas dos movimentos dos jogadores.

Segundo Silva (2022), utilizou-se a biblioteca OpenCV junto a modelos de Deep Neural Networks (DNN) para efetuar a detecção de poses e a identificação precisa da bola. Tal combinação permitiu a exploração da relação entre a postura do jogador e o momento de contato com a bola, estabelecendo a base para a criação de um modelo preditivo. Silva (2022) também afirma que o modelo foi capaz de inferir a direção provável do chute com base na dinâmica e na orientação corporal do jogador momentos antes do chute, demonstrando a potencial da visão computacional na análise esportiva. A Figura 2 mostra a detecção do jogador e da bola.

Figura 2 – Jogador e bola detectados

Pessoas em campo de futebol

Descrição gerada automaticamente

Fonte: Silva (2022).

De acordo com Silva (2022), a coleta de dados foi uma das principais etapas do sistema. Nelas, realizaram-se gravações tanto diurnas quanto noturnas em um campo de futebol padrão, empregando múltiplas câmeras de alta definição posicionadas em ângulos diferentes para capturar todos os movimentos dos jogadores durante as cobranças de pênalti. Ainda segundo Silva (2022), este arranjo assegurou a captura de uma ampla quantidade de dados sob variadas condições de iluminação, refletindo as situações reais enfrentadas em partidas oficiais.

Para testar a eficácia do sistema, um grupo de 50 jogadores, divididos igualmente entre profissionais e amadores, realizaram cobranças de pênalti em grupos de dez tentativas cada. De acordo com Silva (2022), durante essas sessões, o sistema foi encarregado de identificar o momento exato do contato do pé com a bola e de analisar a postura corporal dos jogadores imediatamente antes do chute. Os resultados demonstraram uma precisão de 94% na identificação correta do momento do chute para jogadores profissionais e de 89% para amadores. Além disso, a análise da postura revelou que, em 85% dos casos, era possível prever a direção do chute com base na orientação dos ombros e quadris do jogador momentos antes da execução da penalidade.

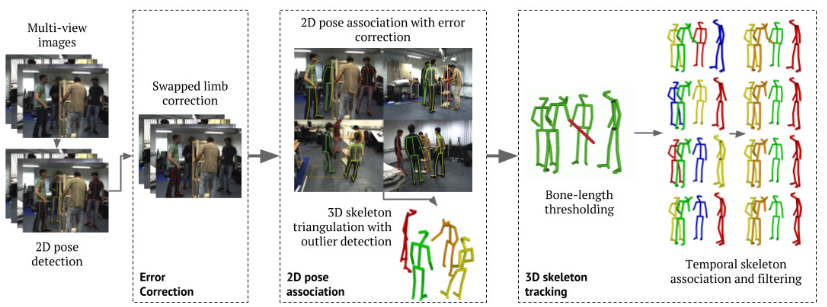
Segundo Silva (2022), os testes também exploraram a solidez do algoritmo em diferentes condições de iluminação, notando-se uma diminuição na precisão da detecção sob iluminação noturna para 92% para profissionais e 86% para amadores, demonstrando a importância de ajustes no algoritmo para maximizar o desempenho em todas as condições ambientais. Embora promissor, Silva (2022), reconhece a necessidade novas pesquisas para abordar outras variáveis cinemáticas e biomecânicas.

## Multi-Person 3D Pose Estimation and Tracking in Sports

O estudo realizado por Bridgeman *et al*. (2019) visava enfrentar os desafios relacionados à estimativa e rastreamento da pose tridimensional de múltiplas pessoas em ambientes esportivos dinâmicos. Buscando entender a complexidade apresentada por esses cenários, que incluem movimentos rápidos dos atletas, interações físicas frequentes entre jogadores, e problemas comuns de visibilidade como obstrução substanciais e falhas de calibração das câmeras, os autores propuseram uma solução que busca não apenas aumentar a precisão da estimativa de pose, mas também acelerar o processamento em tempo real. A partir disso, Bridgeman *et al*. (2019) desenvolveram um método para gerar e rastrear esqueletos 3D utilizando informações de múltiplas câmeras sem depender de modelos baseados em aparência, que são menos eficazes em ambientes esportivos devido às semelhanças de vestimentas e rápida mudança de cenários.

O método proposto por Bridgeman *et al*. (2019) consiste inicialmente na detecção independente de poses 2D em diferentes ângulos de captura das imagens (multivisão). Posteriormente, os erros de saída do detector de poses são corrigidos antes de aplicar um algoritmo guloso (*greedy*) para associar as detecções de poses 2D. A partir dessas junções, esqueletos 3D são gerados utilizando triangulação para gerar os esqueletos 3D, ao qual permitem o rastreamento ao longo do tempo. A Figura 3 exibe as etapas da representação 3D dos esqueletos gerados a partir de imagens 2D

Figura 3 – Etapas da representação 3D dos esqueletos



Fonte: Bridgeman *et al*. (2019).

Bridgeman *et al*. (2019) desenvolveram técnicas para corrigir erros comuns associados aos detectores de pose multivisão, que são particularmente desafiadores em ambientes esportivos. Os autores focaram em problemas como (i) poses divididas, onde uma única pessoa é detectada como várias entidades separadas, (ii) poses fundidas, onde partes de múltiplos indivíduos são erroneamente combinadas em uma única detecção, e (iii) junções trocadas, onde as conexões entre as partes do corpo são atribuídas incorretamente. Para resolver essas questões, Bridgeman *et al*. (2019) utilizaram informações de várias imagens (múltiplas visões) para identificar e corrigir esses erros de forma eficaz. Segundo os autores, essas correções são cruciais para garantir a precisão na geração e no rastreamento subsequente de esqueletos 3D, tornando o método robusto mesmo em situações aos quais as detecções possuem ruídos ou são incompletas devido à rápida ação e interações complexas entre os jogadores.

Para validar essa abordagem, Bridgeman *et al*. (2019) utilizaram uma série de conjuntos de dados esportivos, conhecidos por sua má calibração e com frequentes erros de detecção de pose. Esses conjuntos de dados representam cenários dinâmicos, como jogos de futebol e basquete, onde os participantes se movem rapidamente e interagem fisicamente de maneiras que podem confundir os algoritmos de detecção. Segundo os autores, está abordagem mostrou-se capaz de superar os métodos existentes, demonstrando uma melhoria substancial na velocidade de processamento, alcançando mais de 110 quadros por segundo no rastreamento de esqueletos 3D no conjunto de dados "Shelf". Este conjunto é particularmente complexo devido à alta densidade de jogadores e à dinâmica variável, que frequentemente inclui oclusões e rápida mudança de posições, desafiando os limites do que era anteriormente possível em termos de análise de movimento esportivo em tempo real.

Bridgeman *et al*. (2019) também criaram conjuntos de imagens com diferentes números de pessoas e câmeras, e com arranjos de câmeras de linha de base estreita e ampla. Neste caso, as pessoas estão contidas dentro de um círculo de raio de 2,5m, e as câmeras em um círculo de raio de 5m. As câmeras de linha de base estreita têm um espaçamento de 10 graus, e as câmeras de linha de base ampla estão equidistantes ao redor do círculo. Segundo os autores, ao executar o detector de pose em cada conjunto de imagens, gerando uma matriz binária onde cada célula representa um par de poses; 1 representa uma correspondência e 0 caso contrário. Bridgeman *et al*. (2019) relatam que o método alcança uma precisão mais alta com o arranjo de câmera de linha de base ampla para um número menor de câmeras e um maior número de pessoas. Isso se deve à melhoria da robustez à oclusão e à precisão da triangulação com visualizações de linha de base ampla. Notavelmente, segundo os autores, o método alcança 100% em cenas com duas pessoas para todas as configurações de câmera.

Por fim, Bridgeman *et al*. (2019) concluem que o método representa um avanço significativo na análise esportiva, proporcionando uma ferramenta poderosa para estimação e rastreamento de pose 3D em tempo real. Entretanto, os autores reconhecem que o método ainda enfrenta desafios com oclusões extremas e detecções errôneas. Além disso, também sugerem a utilização de outros algoritmos para realizar a correção de erros e associação de poses.

# proposta

Nesta seção serão apresentadas as justificativas para a concepção deste projeto, também serão apresentados os Requisitos Funcionas (RF) juntamente com os Requisitos Não Funcionais (RNF) e as metodologias que serão utilizadas.

## JUSTIFICATIVA

No Quadro 1 será apresentado o comparativo entre os trabalhos correlatos no entorno das características destacadas de cada um deles, com ele consegue-se entender melhor os objetivos que serão desenvolvidos neste projeto.

Quadro 1 – Comparativo dos trabalhos correlatos

|  |  |  |  |
| --- | --- | --- | --- |
| Trabalhos Correlatos  Características | Pinheiro (2022) | Silva (2022) | Bridgeman *et al*. (2019) Bridgeman (2019) |
| Objetivo | Avaliação de estratégias de cobranças de pênalti em competições de futebol usando OpenPose e análise notacional | Identificação do momento exato do contato do pé do jogador com a bola durante a cobrança de pênalti | Estimativa e rastreamento da pose 3D de múltiplas pessoas em ambientes esportivos dinâmicos |
| Tecnologias utilizadas | OpenPose, Kinovea | OpenCV, Deep Neural Networks DNN) | Detecção independente de poses 2D, algoritmo guloso para associar poses, triangulação para gerar esqueletos 3D |
| Métodos de coleta de dados | Análise de vídeos de cobranças de pênalti, cálculo de velocidade da bola | Gravação de vídeos em campo de futebol, uso de múltiplas câmeras de alta definição | Gravação de vídeos esportivos, criação de conjuntos de imagens com diferentes configurações de câmera e pessoas |
| Base de dados utilizadas | 34 vídeos de cobranças de pênaltis das principais ligas europeias e da UEFA Champions League (2017-2020). | 120 filmagens de cobranças de pênalti de um campo society, com diferentes ângulos e iluminação, 60 fps. | Datasets "Campus" e "Shelf". |
| Métricas de avaliação | Precisão na identificação da direção do chute, antecipação dos movimentos dos goleiros | Precisão na identificação do momento do chute, precisão na previsão da direção do chute | Precisão na estimativa e rastreamento da pose 3D, velocidade de processamento, robustez à oclusão e detecção de erros |

Fonte: elaborado pelo autor.

A partir do Quadro 1 pode-se observar que os três trabalhos abordam a análise de cobranças de pênalti e estimativa de pose em ambientes esportivos, cada um com seus objetivos e métodos específicos. Pinheiro (2022) se concentra na avaliação das estratégias de cobranças de pênalti, utilizando o OpenPose e análise notacional para compreender a relação entre a postura dos jogadores e o sucesso nas cobranças. Enquanto Silva (2022) busca identificar o momento exato do contato do pé do jogador com a bola durante a cobrança, empregando a biblioteca OpenCV e modelos de Deep Neural Networks (DNN). Por outro lado, Bridgeman *et al*. (2019) tinham como objetivo principal a estimativa e rastreamento da pose 3D de múltiplas pessoas em ambientes esportivos dinâmicos, utilizando detecção independente de poses 2D e triangulação para gerar esqueletos 3D.

Em relação aos métodos de coleta de dados, Pinheiro (2022) analisa vídeos de cobranças de pênalti de competições de futebol de alto nível, Silva (2022) grava vídeos em campo de futebol, empregando múltiplas câmeras de alta definição. Já Bridgeman *et al*. (2019) gravam vídeos esportivos e criam conjuntos de imagens com diferentes configurações de câmera e pessoas, utilizando conjuntos de dados específicos.

Quanto a abordagem tecnológica, Pinheiro (2022) optou pela utilização do OpenPose, uma solução avançada de visão computacional, em conjunto com o software Kinovea, amplamente reconhecido por sua precisão na análise de movimentos. Por outro lado, Silva (2022) emprega a biblioteca OpenCV em combinação com modelos de Deep Neural Networks (DNN), destacando-se pela capacidade de processamento complexo de dados e reconhecimento de padrões. Em contrapartida, Bridgeman *et al*. (2019) adotam uma abordagem diferente, fazendo uso de detecção independente de poses 2D e triangulação para a geração de esqueletos 3D, um método que prioriza a robustez e a eficiência na estimativa de poses tridimensionais em ambientes esportivos dinâmicos

No que diz respeito as métricas de avaliação, Pinheiro (2022) se concentra na avaliação da precisão na identificação da direção do chute e na antecipação dos movimentos dos goleiros, visando compreender a eficácia das estratégias adotadas pelos jogadores. Já Silva (2022) direciona sua análise para medir a precisão na identificação do momento exato do chute e na previsão da direção do mesmo, destacando a importância de capturar e interpretar os movimentos cruciais durante a cobrança. Por sua vez, Bridgeman *et al*. (2019) realizam uma abordagem abrangente, incluindo a análise da precisão na estimativa e rastreamento da pose 3D, a velocidade de processamento do algoritmo, além da robustez frente à oclusão e à detecção de erros, oferecendo uma visão detalhada do desempenho do método em cenários dinâmicos e desafiadores.

Considerando as abordagens dos trabalhos correlatos, este trabalho torna-se relevante pois visa desenvolver um artefato computacional que seja capaz de prever e mostrar, em tempo real, a direção dos chutes durante as cobranças de pênalti. Dentre as técnicas a serem exploradas estão algoritmos de aprendizado de máquina, como Convolutional Neural Network (CNNs) e métodos de visão computacional, como a detecção de pontos de referência anatômicos, análise de movimento, detecção e rastreamento de poses. Essas abordagens prometem não só melhorar o desempenho individual dos goleiros ou sua habilidade de antecipação, mas também proporcionar uma compreensão mais abrangente das dinâmicas do jogo, aumentando a precisão das previsões e fornecendo informações importantes que podem ser utilizados para aprimorar as estratégias de treinamento e desempenho. Com isso, espera-se não apenas uma melhoria no desempenho dos atletas ao bater ou defender pênaltis, mas também um aumento na interpretação e competitividade do jogo, contribuindo para avanços significativos no campo da análise esportiva, fornecendo detalhes sobre a mecânica e a tomada de decisão dos jogadores em campo.

## REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

Os requisitos estão divididos em duas partes, uma relacionada a aplicação que permite o cadastro de vídeos e a visualização dos dados em tempo real por treinadores e goleiros, e outra relacionada ao método de detecção, rastreamento e estimativa das poses dos jogadores antes da cobrança de pênaltis.

A aplicação web deverá:

1. permitir que o usuário cadastre de vídeos de diferentes ângulos dos jogadores de futebol realizando cobranças de pênaltis (Requisito Funcional - RF);
2. permitir que o usuário faça o *upload* dos vídeos que serão usados para detectar, rastrear e estimar as poses dos jogadores (RF);
3. proporcionar *feedback* e análises pós-treino para goleiros e treinadores revisarem o desempenho (RF);
4. exibir as previsões de direção de chutes em uma interface web de usuário intuitiva para os goleiros e treinadores (RF);
5. utilizar a linguagem de programação Python juntamente com o *framework* Flask para o desenvolvimento da API rest (Requisito Não Funcional - RNF).

O método de detecção, rastreamento e estimativa das poses dos jogadores antes da cobrança de pênaltis deverá:

1. utilizar técnicas de aprendizado de máquina para a detecção das poses dos jogadores tais como redes neurais convolucionais (YOLO, Inception-v3, Xception; DensePose, PoseNet e OpenPose) (RF);
2. utilizar técnicas de visão computação para estimar e prever a direção dos chutes (RF);
3. identificar e interpretar padrões de movimento correlacionados à direção dos chutes (RF);
4. armazenar dados de treino e jogos para aprimoramento contínuo do algoritmo de previsão (RF);
5. ser desenvolvido na linguagem de programação Python (RNF);
6. utilizar o Keras para o auxílio na utilização de um algoritmo de detecção de objetos e o OpenCV para auxiliar na utilização de um algoritmo de rastreamento e estimativa da posição e chutes (RNF);
7. analisar em tempo real a postura dos jogadores durante as cobranças de pênalti (RNF).

## METODOLOGIA

O trabalho será desenvolvido observando as seguintes etapas:

1. levantamento bibliográfico: realizar o levantamento bibliográfico sobre estimação de pose humana, redes neurais convolucionais (técnicas para rastreamento, detecção de objetos e estimação da pose) e trabalhos correlatos;
2. validação dos requisitos: levando em consideração os dados da etapa anterior, reavaliar os requisitos para o modelo proposto;
3. coleta de vídeos de cobranças de pênaltis: coletar vídeos que serão usadas nos testes do método proposto. Obter vídeos de cobranças de pênaltis aos quais caracterizem os movimentos dos jogadores que precedem o chute;
4. pesquisa e escolha do algoritmo de detecção de objetivos: pesquisar os principais algoritmos de detecção de objetos, escolhendo o mais adequado para identificação do jogador e da bola;
5. pesquisa e escolha do algoritmo de estimação e rastreamento das poses: pesquisar os principais algoritmos de estimação e rastreamento poses dos jogadores, escolhendo o mais adequado para o desenvolvimento do trabalho;
6. desenvolvimento do modelo: desenvolver o método de detecção e rastreamento das poses dos jogadores e a extração de características para prever a direção do chute com base na pose do jogador momentos antes de executar o chute. Para isso, será utilizada a linguagem de programação Python e a biblioteca Keras juntamente com a biblioteca OpenCV, redes neurais convolucionais tais como, YOLO, Inception-v3, Xception; DensePose, PoseNet e OpenPose;
7. validação e testes do modelo: em paralelo ao desenvolvimento, verificar a assertividade do método proposto a partir da detecção e estimativa correta da pose do jogador, assim como sua capacidade de prever a direção dos chutes. Avaliar o impacto da tecnologia no treinamento e desempenho dos goleiros, e refinar o modelo com base nos feedbacks dos usuários e resultados dos testes;
8. elicitação de requisitos da interface gráfica com o usuário: detalhar e reavaliar os requisitos propostos para a aplicação e, se necessário, especificar outros a partir das necessidades observadas ao longo do trabalho;
9. especificação: elaborar os diagramas de casos de uso e de classes de acordo com a Unified Modeling Language (UML), utilizando a ferramenta Astah;
10. desenvolvimento: a partir do item (i) desenvolver a aplicação que permita aos treinadores e goleiros visualizarem e interpretarem os dados em tempo real, melhorando o processo de tomada de decisão durante os jogos e treinamentos. Para isso, será utilizado a linguagem de programação Python com o *framework* Flask;
11. testes da aplicação: elaborar testes para validar a usabilidade da aplicação.

As etapas serão realizadas nos períodos relacionados no Quadro 2.

Quadro 2 – Cronograma de atividades a serem desenvolvidas

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2024 | | 2025 | | | | | | | | | | | |
|  | dez. | | jan. | | fev. | | mar. | | abr. | | maio | | jun. | |
| etapas / quinzenas | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| levantamento bibliográfico |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| validação dos requisitos |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| coleta de vídeos de cobranças de pênaltis |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| pesquisa e escolha do algoritmo de detecção de objetivos |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| pesquisa e escolha do algoritmo de estimação e rastreamento das poses |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| desenvolvimento do modelo |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| validação e testes do modelo |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| elicitação de requisitos da interface gráfica com o usuário |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| especificação |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| desenvolvimento |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| testes da aplicação |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Fonte: elaborado pelo autor.

# REVISÃO BIBLIOGRÁFICA

Nesta seção serão apresentados os principais conceitos que fundamentam o estudo proposto. A subseção 4.1 aborda estimação de pose humana, uma técnica crucial para a identificação das articulações do corpo humano em imagens ou vídeos, com foco especial em suas aplicações no contexto esportivo. Na subseção 4.2, conceitua-se o aprendizado de máquina (AM), considerando suas categorias, técnicas e aplicações na análise de imagens.

## ESTIMAÇÃO DE POSE HUMANA

A estimativa de pose humana refere-se ao processo de identificar a localização das articulações do corpo humano em imagens ou vídeos. Essa área é fundamental para diversas aplicações, como reconhecimento de atividades, interação homem-máquina, realidade aumentada e animação. A precisão e a eficiência dessas estimativas têm sido significativamente aprimoradas com o desenvolvimento de tecnologias avançadas e técnicas de visão computacional, permitindo a captura de contextos mais complexos (GÜLER *et al*., 2018).

Segundo Cao *et al*. (2017), no processo de estimativa de pose, é crucial considerar várias variáveis que influenciam a precisão. Entre elas estão a postura, a orientação, as oclusões e as variações nas condições de iluminação e vestimenta. Essas variáveis são essenciais para garantir a precisão das estimativas, especialmente em ambientes dinâmicos e imprevisíveis como os esportivos. A postura e a orientação do corpo, por exemplo, influenciam diretamente a forma como as articulações se movem e como são capturadas pelas câmeras. Já as oclusões, que ocorrem quando partes do corpo são ocultadas por outras partes ou por objetos no ambiente, podem comprometer a precisão da estimativa, exigindo técnicas avançadas para resolver essas ambiguidades. Além disso, variações nas condições de iluminação e na vestimenta podem alterar a aparência visual das articulações, dificultando sua detecção precisa (CHEN *et al*., 2017).

Güler *et al*. (2018) ressaltam que essas estimativas são particularmente importantes na análise de movimentos esportivos, onde a capacidade de monitorar desempenhos e prever movimentos futuros é crucial. Segundo os autores, a identificação precisa da posição das articulações permite uma análise detalhada da biomecânica dos atletas, auxiliando na melhoria do desempenho e na prevenção de lesões. Além disso, a estimativa de pose humana tem aplicações em áreas como a robótica, onde o entendimento dos movimentos do corpo humano é essencial para a interação natural entre homem e máquina.

De acordo com Mehta *et al*. (2020), um dos desafios quanto a estimativa de pose é lidar com a grande variabilidade de posturas e movimentos humanos. A amplitude de possibilidades de movimentação torna a tarefa de identificar precisamente as articulações extremamente complexa. Essa complexidade aumenta ainda mais em contextos específicos, como os esportes, onde os atletas realizam movimentos complexos e imprevisíveis. Os autores ainda destacam que a análise da marcha é outra aplicação importante, envolvendo o estudo detalhado dos movimentos das pernas e dos pés durante a caminhada ou corrida. Este tipo de análise é vital para identificar padrões de movimento que possam indicar problemas como pronação excessiva ou falta de flexibilidade nos tornozelos. Neste caso, técnicas de estimativa de pose humana ajudam a capturar esses detalhes, permitindo uma avaliação precisa que pode ser usada para personalizar programas de reabilitação ou treinamento. Para Cao *et al*. (2017), essas técnicas permitem mapear cada articulação e acompanhar seus movimentos ao longo do tempo, o que é particularmente útil em esportes, onde a precisão dos movimentos pode ser a diferença entre sucesso e falha.

Segundo Chen *et al*. (2017), além das aplicações em análise de movimentos, a estimativa de pose humana é amplamente utilizada na área esportiva para monitorar a reabilitação e recuperação de atletas lesionados. Durante a reabilitação, essas técnicas podem ser usadas para monitorar a progressão do atleta, avaliando como as articulações e músculos estão se recuperando de uma lesão. A análise da marcha e da biomecânica do movimento beneficia-se significativamente dessas técnicas, permitindo a avaliação detalhada da recuperação dos atletas e ajudando a ajustar os programas de treinamento de forma personalizada. Por exemplo, após uma lesão no joelho, a análise da marcha pode identificar se o atleta está retornando à sua forma normal de caminhar ou se ainda está compensando a lesão com a outra perna, o que pode levar a novas lesões. Neste sentido, a avaliação contínua permite que os treinadores e fisioterapeutas ajustem os métodos de reabilitação em um tempo muito menor, garantindo que o atleta evolua de maneira mais rápida, segura e eficaz. Isso não só melhora a eficácia dos treinos, mas também acelera a recuperação, permitindo que o indivíduo volte ao seu desempenho máximo o mais rápido possível (CAO *et al*., 2017).

Por fim, de acordo com Güler *et al*. (2018), a estimativa de pose também enfrenta desafios relacionados à eficiência computacional. Para muitas aplicações, como sistemas de tempo real em robótica e realidade aumentada, é essencial que o processo de estimativa de pose seja realizado de forma rápida e com baixo consumo de recursos computacionais. Nesse sentido, a pesquisas têm se concentrado no desenvolvimento de modelos de aprendizado de máquina para realizar a estima da pose humana.

## Aprendizado de maquina

O Aprendizado de Máquina (AM) tem se tornado essencial no desenvolvimento de sistemas inteligentes, especialmente na análise de imagens. Conforme destacado por Janiesch *et al*. (2021), AM refere-se à capacidade dos sistemas de aprenderem a partir de dados de treinamento específicos, automatizando a construção de modelos analíticos e resolvendo tarefas associadas. No contexto da visão computacional, o aprendizado profundo, uma subcategoria de AM, destaca-se por superar modelos tradicionais em várias aplicações, incluindo a detecção e análise de imagens (GOODFELLOW *et al*., 2016).

Segundo Jordan e Mitchell (2015), as técnicas de Aprendizado de Máquina podem ser classificadas em três categorias principais: aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço. O aprendizado supervisionado envolve o treinamento de modelos com conjuntos de dados rotulados, permitindo que o sistema aprenda a mapear entradas para saídas conhecidas. Por outro lado, o aprendizado não supervisionado foca na identificação de padrões ocultos em dados não rotulados, enquanto o aprendizado por reforço se baseia em interações com o ambiente, onde o modelo recebe recompensas ou penalidades conforme suas ações (JORDAN; MITCHELL, 2015).

De acordo com Janiesch *et al*. (2021), a construção de modelos analíticos automatizados com aprendizado profundo envolve várias etapas, como entrada de dados, extração de características, construção do modelo e avaliação do modelo. No caso da análise de imagens, a entrada de dados pode incluir fotos e vídeos, que são processados para extrair informações relevantes (JANIESCH *et al*., 2021). A extração de características é uma etapa crucial, sendo as *Convolutional Neural Networks* (CNNs) altamente eficazes para essa tarefa. As CNNs processam dados visuais usando camadas convulsionais que extraem automaticamente características importantes, como bordas, texturas e formas (O'SHEA; NASH, 2015).

Ainda segundo O'Shea e Nash (2015), ferramentas baseadas em aprendizado de máquina têm sido fundamentais para diversas aplicações na análise de imagens, incluindo a detecção de pontos-chave do corpo em tempo real. Essas ferramentas fornecem dados que são utilizados para melhorar o desempenho dos atletas e desenvolver estratégias de jogo em diferentes esportes e modalidades. Um exemplo notável é o OpenPose, que permite a detecção precisa de pontos-chave do corpo, auxiliando na análise de movimentos e na criação de programas de treinamento personalizados (CAO et al., 2017). O uso de redes neurais profundas tem proporcionado avanços significativos na área, com métodos como o DeepPose, que utiliza uma cascata de regressões para melhorar a precisão das estimativas (TOSHEV; SZEGENY, 2014).

De acordo com Toshev e Szegeny (2014), outra contribuição significativa é o DensePose, que se concentra na estimativa densa das superfícies do corpo humano, mapeando cada pixel da imagem para a superfície 3D correspondente do corpo humano. Segundo os autores, este método oferece uma supervisão detalhada e precisa, essencial para aplicações que requerem um entendimento granular do movimento corporal, como a análise de técnicas de atletas e a melhoria de performance em esportes competitivos.

Janiesch *et al*. (2021) ressaltam que arquiteturas avançadas como *You Only Look Once* (YOLO), Inception-v3, Xception, DensePose, PoseNet e OpenPose são amplamente utilizadas para a detecção e análise de imagens. O YOLO é conhecido por sua velocidade e precisão na detecção de objetos em tempo real, tornando-se ideal para aplicações que exigem processamento rápido (REDMON *et al*., 2016). O Inception-v3 e o Xception são redes profundas que oferecem alta precisão na classificação de imagens, utilizando módulos convulsionais eficientes para melhorar o desempenho e reduzir a complexidade computacional (SZEGEDY *et al*., 2016; CHOLLET, 2017). O PoseNet é utilizado especificamente para a estimação de pose, permitindo a detecção de poses humanas em diferentes cenários com alta precisão (KENDALL; CIPOLLA, 2015).

Segundo Janiesch *et al*. (2021), a avaliação da qualidade dos modelos envolve métricas de desempenho como precisão, recall e F1-score, além de aspectos computacionais e de interpretabilidade. Ainda de acordo com os autores, esses modelos de aprendizado supervisionado frequentemente utilizam validação cruzada k-fold para prevenir *overfitting* e determinar o desempenho em dados não vistos. Além disso, medidas sensíveis ao custo, como o custo médio por observação prevista, são úteis em cenários onde erros de previsão estão associados a estruturas de custo assimétricas.

Referências

BRIDGEMAN, S. et al. Multi-Person 3D Pose Estimation and Tracking in Sports. **In**: CVPR Workshops. 2019. Disponível em: https://openaccess.thecvf.com/content\_CVPRW\_2019/html/CVSports/Bridgeman\_Multi-Person\_3D\_Pose\_Estimation\_and\_Tracking\_in\_Sports\_CVPRW\_2019\_paper.html.

CAO, Z. et al. **Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition** (CVPR), 2017. Disponível em: https://arxiv.org/abs/1611.08050.

CHEN, X. *et al*. **Pose Analysis and Biomechanics in Sports. Journal of Sports Science**, 2017. Disponível em: https://www.dartfish.com/technology.

CHOLLET, F. **Xception**: Deep Learning with Depthwise Separable Convolutions. **In**: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. Disponível em: https://ieeexplore.ieee.org/document/8099678.

CUI, R. et al. **Learning to Refine Human Pose Estimation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, 2018 Disponível em: https://arxiv.org/abs/1804.07909.

CURLEY, J. P.; ROEDER, O. **English Soccer’s Mysterious Worldwide Popularity**. Contexts, Vol. 15, No. 1, pp. 78-81, 2016. DOI: 10.1177/1536504216628849.

DARTFISH. **Dartfish technology**. 1999. Disponível em: https://www.dartfish.com/technology.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep learning. MIT Press, 2016.

GÜLER, R. A. et al.  **DensePose: Dense Human Pose Estimation In The Wild. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition** (CVPR), 2018. Disponível em: https://arxiv.org/abs/1802.00434.

JANIESCH, C. et al. **Advanced machine learning in business. Business & Information Systems Engineering**, 2021. Disponível em: https://link.springer.com/article/10.1007/s12525-021-00475-2.

JORDAN, M. I.; MITCHELL, T. M. **Machine learning: Trends perspectives and prospects. Science**, 2015. Disponível em: https://doi.org/10.1126/science.aaa8415.

KENDALL, A.; CIPOLLA, R. Geometric Loss Functions for Camera Pose Regression with Deep Learning. **In**: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015. Disponível em: https://arxiv.org/abs/1704.00390

KLEINA, T. et al. Artificial Intelligence in Football: How ANN and SVM Can Predict Match Outcomes. **In**: Journal of Sports Analytics, 2021. Disponível em: https://www.scielo.br/j/jpe/a/Z3PVmqLxFcCn68Ns7SG87Bx/?format=pdf&lang=en.

MEMAR, S. et al.. **Deep Learning-Based Human Pose Estimation: A Survey. IEEE Transactions on Human-Machine Systems**, 50(4), 350-360, 2020. Disponível em: https://arxiv.org/abs/2012.13392.

MEHTA, D. et al. **Deep Convolutional Networks for High-Resolution 3D Human Pose Estimation with Global 2D Alignment. International Journal of Computer Vision** (IJCV), 128(4), 930-956, 2020. Disponível em: https://arxiv.org/abs/1705.05329.

O'SHEA, K.; NASH, R**. An Introduction to Convolutional Neural Networks**. 2015. Disponível em: https://arxiv.org/abs/1511.08458.

PINHEIRO, J**. Estimação de Pose Corporal Integrada à Análise Notacional: Uma Nova Abordagem para Analisar Estratégias de Chutes de Pênalti no Futebol de Elite**. 2022. Disponível em: https://www.frontiersin.org/articles/10.3389/fspor.2022.818556/full.

PROZONE. **Prozone**. 1995. Disponível em: http://prozonesports.stats.com/.

REDMON, J. *et al.* You Only Look Once: Unified, Real-Time Object Detection. **In**: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016. Disponível em: https://ieeexplore.ieee.org/document/7780460.

SILVA, F. M. **Explorando a captação de movimento dos jogadores de futebol em cobranças de pênaltis**. São Paulo: Universidade Presbiteriana Mackenzie, 2022. Disponível em: https://adelpha-api.mackenzie.br/server/api/core/bitstreams/df05baa1-f852-4de9-9aee-e6c7ef22ad5b/content.

SZEGEDY, C. *et al*. Rethinking the Inception Architecture for Computer Vision. **In**: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016. Disponível em: https://ieeexplore.ieee.org/document/7780677.

TOSHEV, A.; SZEGENY, C. DeepPose: **Human Pose Estimation via Deep Neural Networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**, 2014. Disponível em: https://ieeexplore.ieee.org/document/6909610.

ZHANG, B. et al. **Artificial Intelligence-Based Joint Movement Estimation Method for Football Players in Sports Training. Mobile Information Systems**, Volume 2021, Article ID 9956482, 9 pages. Disponível em: https://doi.org/10.1155/2021/9956482.

FORMULÁRIO DE avaliação BCC – PROFESSOR TCC I – projeto

Avaliador(a): Dalton Solano dos Reis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ASPECTOS AVALIADOS | | atende | atende parcialmente | não atende |
| ASPECTOS TÉCNICOS | 1. INTRODUÇÃO   O tema de pesquisa está devidamente contextualizado/delimitado? | X |  |  |
| O problema está claramente formulado? | X |  |  |
| 1. OBJETIVOS   O objetivo principal está claramente definido e é passível de ser alcançado? | X |  |  |
| Os objetivos específicos são coerentes com o objetivo principal? | X |  |  |
| 1. JUSTIFICATIVA   São apresentados argumentos científicos, técnicos ou metodológicos que justificam a proposta? | X |  |  |
| São apresentadas as contribuições teóricas, práticas ou sociais que justificam a proposta? | X |  |  |
| 1. METODOLOGIA   Foram relacionadas todas as etapas necessárias para o desenvolvimento do TCC? | X |  |  |
| Os métodos, recursos e o cronograma estão devidamente apresentados? | X |  |  |
| 1. REVISÃO BIBLIOGRÁFICA (atenção para a diferença de conteúdo entre projeto e pré-projeto)   Os assuntos apresentados são suficientes e têm relação com o tema do TCC? | X |  |  |
| ASPECTOS METODOLÓGICOS | 1. LINGUAGEM USADA (redação)   O texto completo é coerente e redigido corretamente em língua portuguesa, usando linguagem formal/científica? | X |  |  |
| A exposição do assunto é ordenada (as ideias estão bem encadeadas e a linguagem utilizada é clara)? | X |  |  |
| 1. ORGANIZAÇÃO E APRESENTAÇÃO GRÁFICA DO TEXTO   A organização e apresentação dos capítulos, seções, subseções e parágrafos estão de acordo com o modelo estabelecido? | X |  |  |
| 1. ILUSTRAÇÕES (figuras, quadros, tabelas)   As ilustrações são legíveis e obedecem às normas da ABNT? | X |  |  |
| 1. REFERÊNCIAS E CITAÇÕES   As referências obedecem às normas da ABNT? | X |  |  |
| As citações obedecem às normas da ABNT? | X |  |  |
| Todos os documentos citados foram referenciados e vice-versa, isto é, as citações e referências são consistentes? | X |  |  |

|  |  |  |
| --- | --- | --- |
| O projeto de TCC será reprovado se:   * qualquer um dos itens tiver resposta NÃO ATENDE; * pelo menos **4 (quatro)** itens dos **ASPECTOS TÉCNICOS** tiverem resposta ATENDE PARCIALMENTE; ou * pelo menos **4 (quatro)** itens dos **ASPECTOS METODOLÓGICOS** tiverem resposta ATENDE PARCIALMENTE. | | |
| **PARECER**: | ( X ) APROVADO | ( ) REPROVADO |