

# Utilização de Técnicas de Visão Computacional para Detecção de Impedimento no Futebol

Daniel Araujo Cavassani, Felipe Keller Baltor e Othon Daiki Ishiyi

## I. INTRODUÇÃO

A análise de jogos de futebol, um esporte que captura o entusiasmo de milhões em todo o mundo, tem avançado significativamente com o advento de técnicas modernas de visão computacional. Este trabalho foca na utilização dessas técnicas para abordar uma das questões mais desafiadoras e controversas do futebol: a detecção de impedimento. Tradicionalmente dependente da percepção humana e, portanto, suscetível a erros, a determinação precisa de situações de impedimento pode ser significativamente aprimorada com o auxílio da tecnologia.

A pesquisa aqui apresentada desenvolve um método inovador que combina algoritmos avançados de visão computacional para identificar com precisão a posição dos jogadores em campo. O cerne deste método reside na habilidade de distinguir jogadores de objetos irrelevantes e classificá-los corretamente em suas respectivas equipes. Isso é alcançado através do uso de modelos de detecção de keypoints e técnicas de segmentação de imagens, permitindo uma análise detalhada e confiável das posições dos jogadores durante momentos cruciais do jogo.

Um aspecto crucial do nosso método é a identificação do ponto de fuga em imagens de jogos de futebol. Através da análise das linhas de demarcação do campo, que são normalmente visíveis em qualquer condição de jogo, aplicamos operações morfológicas e a Transformada de Hough para detectar e filtrar linhas relevantes. Este processo é vital para estabelecer o ponto de convergência a partir do qual as linhas de impedimento são traçadas, fornecendo uma base sólida para a determinação de situações de impedimento.

Ao abordar o desafio de detecção de impedimento com esta nova metodologia, esperamos contribuir para a redução de erros de arbitragem e para o aumento da justiça e precisão em jogos de futebol. Este estudo não apenas fornece insights valiosos para árbitros e treinadores, mas também representa um passo significativo na interseção entre esporte e tecnologia, abrindo caminho para futuras inovações no campo da análise esportiva.

## II. FUNDAMENTAÇÃO TEÓRICA

A visão computacional, uma subárea da inteligência artificial, desempenha um papel crucial na análise automática de imagens e vídeos. Esta seção aborda os conceitos fundamentais e as técnicas de visão computacional aplicados ao desafio da detecção de impedimento no futebol.

### A. Detecção de Objetos e Keypoints

A detecção de objetos é uma das tarefas fundamentais da visão computacional, que consiste em localizar e identificar objetos dentro de uma imagem. No contexto do futebol, isso se traduz na capacidade de identificar e localizar jogadores e outros elementos relevantes do campo. A técnica de detecção de keypoints, que envolve a identificação de pontos específicos no corpo dos jogadores, como cabeça, ombros e pés, é essencial para avaliar a posição e a postura dos atletas. Modelos pré-treinados, como o Detectron2 da Facebook AI Research, são frequentemente utilizados para essas tarefas, oferecendo uma base robusta para análises mais específicas.

### B. Segmentação de Imagens

A segmentação de imagens é o processo de dividir uma imagem em várias partes ou regiões, geralmente com o objetivo de simplificar ou alterar a representação da imagem para facilitar sua análise. No futebol, a segmentação é usada para isolar o campo de jogo de outros elementos, como a plateia ou o céu, permitindo uma análise focada apenas nos aspectos relevantes para a detecção de impedimento.

### C. Transformada de Hough e Análise Geométrica

A Transformada de Hough é uma técnica que permite a detecção de formas geométricas simples, como linhas, em imagens. No nosso contexto, ela é utilizada para identificar as linhas de demarcação do campo de futebol. Combinando essas linhas detectadas com análises geométricas avançadas, é possível estabelecer o ponto de fuga e, consequentemente, traçar as linhas de impedimento de maneira precisa.

### D. Algoritmos de Clustering

O *clustering* é uma técnica de aprendizado de máquina usada para agrupar dados semelhantes. No caso da classificação de times, algoritmos de clustering como DBSCAN e modelos de mistura gaussiana são aplicados para agrupar jogadores com base em características como a cor das camisas, permitindo a diferenciação entre os times e outros participantes do jogo, como árbitros e goleiros.

Esta fundamentação teórica fornece as bases para as técnicas e metodologias aplicadas no presente estudo, facilitando a compreensão de como a visão computacional pode ser eficazmente empregada na análise de jogos de futebol e na detecção de situações de impedimento.

### III. METODOLOGIA E IMPLEMENTAÇÃO

#### A. Detecção de Jogadores

Para a tarefa de identificar jogadores impedidos, é crucial identificar com precisão a localização dos jogadores. Além disso, utilizar simplesmente as *bounding boxes* de cada jogador não é muito efetivo, pois a *bounding box* ocupa uma área muito maior que a área jogável do jogador (o seu corpo excluindo os braços). Desse modo, além de detectar a posição dos jogadores, deve-se identificar as principais extremidades da região jogável de cada atleta, que são cabeça, ombros, quadris, joelhos e pés. Com esse intuito, utilizou-se um modelo pré-treinado de detecção de *keypoints* de pessoas do *Detectron2*, uma plataforma *open source* de detecção de objetos desenvolvida pela *Facebook AI Research*. Utilizando o *dataset* com várias imagens de lances de futebol, encontrado no artigo de Panse e Mahabaleshwarkar (2020), foi feita uma tentativa de treinar o modelo do *Detectron2* para ele se especializar na detecção de pessoas à longa distância. Contudo, devido ao pequeno tamanho do *dataset*, assim como devido à presença de inconsistências, como a não-presença dos *keypoints* dos árbitros no *dataset*, não foi possível obter resultados satisfatórios, e optou-se manter o modelo original.

Ainda, por ser utilizado um modelo de detecção de pessoas, o *Detectron2* também identifica pessoas indesejadas fora do campo, por exemplo, torcedores. Assim, antes da detecção de *keypoints* também foi realizada uma segmentação do campo de futebol, de modo a identificar apenas jogadores e, possivelmente, árbitros. Para isso também foi utilizado o modelo pré-treinado de segmentação panótica do *Detectron2*, capaz de identificar campos de futebol. Porém, como ilustra a figura a seguir, é necessário fazer um pós-tratamento com a região detectada.



Fig. 1. Em branco, a região detectada pelo modelo de segmentação panótica, em imagem retirada do jogo Vasco x Bragantino

Primeiramente, percebe-se que os jogadores criam "buracos" na região detectada como campo. Para contornar esse problema, aplica-se um algoritmo de dilatação do campo. Também percebe-se que o modelo de segmentação panótica pode classificar regiões espúrias como sendo campo de futebol. Desse modo, para cada região identificada como campo de futebol, é calculada a sua área e somente a maior área é mantida. Assim, é possível, ao fim, utilizar a imagem modificada como máscara da original e obter somente o campo de futebol, como ilustra a figura a seguir.



Fig. 2. Imagem do jogo Vasco (camisa preta) x Bragantino (camisa branca) tratada para exibir apenas o campo de futebol

#### B. Classificação de Times

Após a identificação dos *keypoints* de cada jogador, outra tarefa essencial é identificar os times de cada uma das instâncias. Para realizar essa classificação, foram utilizados algoritmos de *clustering* da biblioteca *Python scikit-learn*, aproveitando-se do fato de que a cor das camisas de cada time, assim como a dos árbitros e dos goleiros, devem ser bastante diferentes entre si.

O primeiro passo, então, é extrair a cor da camisa de cada jogador, que serão as entradas dos algoritmos de *clustering*. Para isso, foi feita a média das cores na região delimitada pelos *keypoints* correspondentes aos ombros e aos quadris, já que essa região normalmente corresponde à camisa do jogador. Além disso, ao fazer a média, atenua-se a interferência de ruídos como a numeração e a cor de pele do jogador.

Já o *clustering* é dividido em duas partes. Primeiramente aplica-se o algoritmo *DBSCAN* (*Density-Based Spatial Clustering of Applications with Noise*), especializado em identificar outliers. Esse algoritmo, portanto, é utilizado para separar o goleiro e o árbitro dos demais jogadores. Depois utiliza-se o algoritmo de modelo de mistura gaussiana para classificar os demais jogadores em um dos dois times possíveis.

Os *clusters* do *DBSCAN* são feitos com base na distância euclidiana dos pontos de entrada. Dado um ponto de entrada, se houver outro ponto dentro de um raio *eps*, os dois pontos são categorizados de forma igual. Contudo, para cores em formato BGR, a distância euclidiana não representa necessariamente a diferença visual entre as cores. Por isso, foi utilizada uma heurística para modificar a cor de entrada, aumentando a maior componente BGR e diminuindo a menor componente. Assim, cores similares terão a distância euclidiana aproximadas, já que as maiores componentes serão similares. Também foi testada a utilização do formato LAB, porém os resultados obtidos foram piores, comparados com as obtidas com a heurística.

Após detectar os goleiros e os árbitros com o *DBSCAN*, utiliza-se o modelo de mistura gaussiana do *scikit-learn* para classificar os demais jogadores entre as duas categorias possíveis: times 1 e 2. Dessa vez, a representação das cores utilizando o formato LAB se mostrou eficiente, sendo esse o formato das entradas desse modelo de *clustering*. A figura

a seguir exemplifica o resultado da classificação de times.



Fig. 3. Imagem do jogo Portugal (camisa vermelha e identificado por keypoints azuis) contra Polônia (camisa branca e identificado por keypoints verdes). Árbitros e goleiro estão classificados com keypoints vermelhos.

### C. Identificação do Ponto de Fuga

A identificação do ponto de fuga é um componente chave na análise de imagens de futebol, especialmente no contexto de determinar as linhas de impedimento. Um desafio enfrentado neste processo é a variabilidade na visibilidade das texturas do gramado, que frequentemente apresentam padrões paralelos de diferentes tons de verde, que poderiam ser utilizadas como uma métrica válida ao detectar paralelas válidas. No entanto, essas texturas nem sempre são facilmente identificáveis, especialmente em imagens de baixa resolução ou em campos com manutenção deficiente.

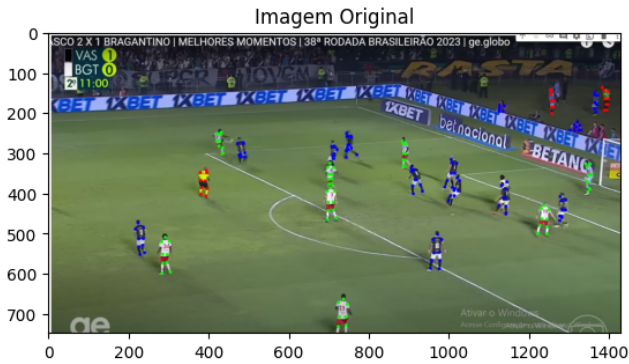


Fig. 4. Imagem original da câmera.

Para superar essas limitações e garantir precisão na análise, a metodologia adotada focou no uso das linhas de demarcação do campo, que são normalmente mais visíveis e consistentes. Isso se torna particularmente útil em lances de impedimento, que ocorrem frequentemente perto da área do gol, onde o acesso a pelo menos duas linhas paralelas de demarcação é mais provável.

O processo começa com a captura da imagem original pela câmera. Posteriormente, essa imagem é convertida para o espaço de cor HSV, facilitando a segmentação de cor para destacar as linhas de demarcação do campo.

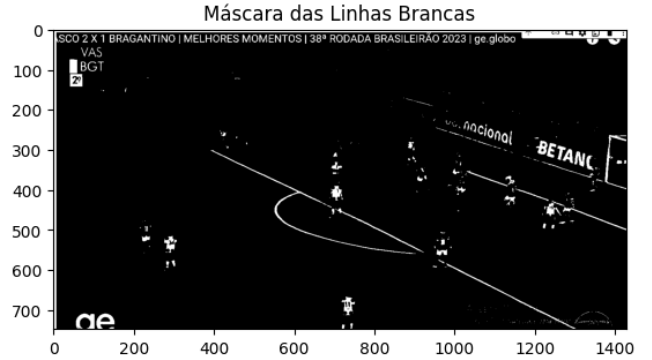


Fig. 5. Imagem no espaço HSV com segmentação de cor.

Em seguida, operações morfológicas são aplicadas para realçar as linhas detectadas, preparando-as para a identificação precisa através da Transformada de Hough. Esta técnica é eficaz para detectar linhas retas na imagem, mas pode gerar um excesso de linhas indesejadas se não for adequadamente controlada.

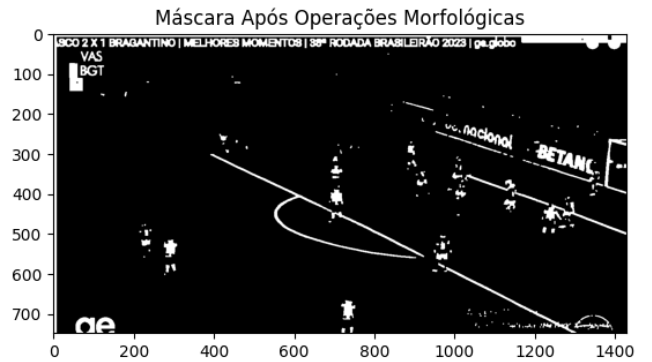


Fig. 6. Resultado após operações morfológicas.

A filtragem das linhas é uma etapa crítica, onde linhas horizontais e verticais irrelevantes são removidas. Este processo foca em manter apenas as linhas úteis para a detecção do ponto de fuga, evitando assim interpretações errôneas.

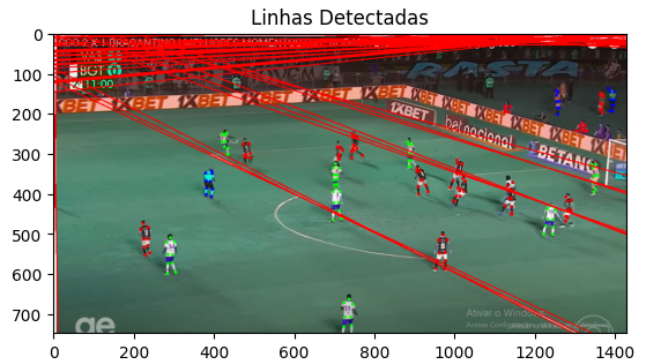


Fig. 7. Detecção de linhas pela Transformada de Hough antes da aplicação dos thresholds.

Após a aplicação dos critérios de filtragem, obtém-se

um conjunto de linhas mais limpo e focado, essencial para determinar com precisão o ponto de fuga na imagem.



Fig. 8. Detecção de linhas pela Transformada de Hough após a aplicação dos thresholds.

Após a filtragem das linhas encontradas pela Transformada de Hough, o próximo passo crucial é a identificação do ponto de convergência. Este ponto é determinante para estabelecer a base a partir da qual as linhas de impedimento serão traçadas. Para encontrar este ponto de convergência, implementou-se um método que considera todas as intersecções possíveis entre as linhas filtradas.

Primeiramente, calcula-se o ponto de intersecção para cada par de linhas utilizando a função `line_intersection`. Esta função resolve um sistema de equações lineares que representa as duas linhas em suas formas paramétricas. Em seguida, com todas as intersecções coletadas, utiliza-se a função `find_convergence_point` para determinar o ponto que minimiza a soma das distâncias quadráticas a todas as intersecções. Esta abordagem, baseada na minimização de uma função de custo, permite encontrar um ponto de convergência mesmo em cenários onde as linhas podem não se cruzar perfeitamente devido a imprecisões na detecção.

O ponto de convergência encontrado é crucial para a próxima etapa: o traçado das linhas de impedimento. Utilizando este ponto como referência, linhas são desenhadas de forma a representar possíveis posições de impedimento no campo.

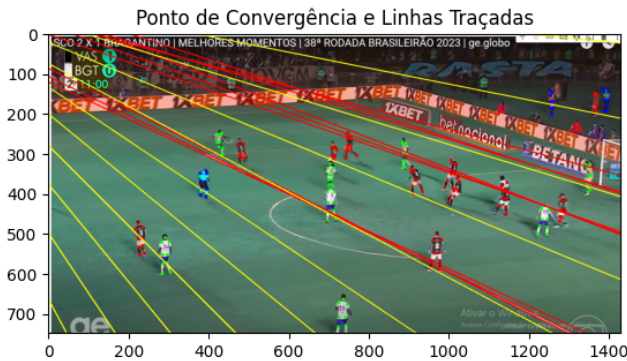


Fig. 9. Linhas de impedimento traçadas a partir do ponto de convergência.

Ao final, a imagem resultante oferece não apenas uma

visualização clara do ponto de convergência, mas também das linhas de impedimento possíveis, fundamentais para a análise de lances de jogo em situações de impedimento.

#### D. Projeção das Partes do Corpo

Dado que a análise do impedimento compreende a avaliação de uma situação tridimensional, existe uma perda de informação quando o método empregado se fia apenas em imagens bidimensionais. Frequentemente a disputa sobre uma situação de impedimento acontece por poucos centímetros de diferença entre um jogador e outro, demandando uma sensibilidade que pode ser abalada pela falta de todos detalhes tridimensionais da situação.

Nesse sentido, a projeção dos keypoints no gramado surge como alternativa de mitigação desse erro. Ao projetar keypoints que representam as partes do corpo de importância para o impedimento estamos eliminando a necessidade de se conhecer a posição tridimensional exata daquela parte e normalizando a análise.

A abordagem desse ponto pode ser feita em vários níveis de sofisticação, partindo de um modelo padrão de tamanho para os jogadores, até a correta determinação da pose dos mesmos e consideração dos tamanhos reais de cada jogador. No presente projeto e como abordagem inicial nós decidimos ignorar esse ponto.

#### E. Linhas de Impedimento

A parte final do modelo é responsável por unir as principais técnicas empregadas (identificação de jogadores e times, bem como seus keypoints, projeção no solo e ponto de fuga) é o traçado da linha de impedimento. Analogamente ao item anterior, buscamos analisar uma situação tridimensional que foi capturada em uma imagem 2D sofrendo efeitos típicos como distorções de perspectiva: as demarcações e entrelinhas do campo, paralelas na realidade, se apresentam convergindo para o ponto de fuga. Dessa maneira, o traçado de linhas de impedimento paralelas se converte no traçado de linhas que passam pelo ponto de fuga.

Assim, todas as linhas unindo keypoints e ponto de fuga são possíveis candidatas à linha de impedimento e a determinação da linha correta pode ser facilmente realizada buscando-se a linha que passa por um jogador que não é goleiro mais próxima à linha que passa pelo goleiro. Daí basta verificar se goleiro e jogador são do mesmo time para inferir o impedimento.

Porém, embora essa abordagem dê conta da maioria dos casos práticos considerando o atual modelo de transmissão de jogos, ainda poderão existir situações atípicas não consideradas nessa lógica.

### IV. RESULTADOS E DISCUSSÃO

#### A. Detecção de Jogadores e Classificação de Times

Para obter métricas que possam avaliar tanto a detecção dos jogadores quanto a classificação deles em times, foi utilizado o *dataset* do artigo de Panse e Mahabaleshwarkar (2020), cujos *keypoints* e classificação de times foram adotados como sendo os valores verdadeiros. Primeiramente, ao



comparar as posições dos jogadores previstas pelo modelo *Detectron2* com os jogadores presentes no *dataset*, foi obtido que 95% dos jogadores do *dataset* também foram identificados pela rede neural. A principal causa de não-detecção foi a sobreposição de dois jogadores na imagem. Dessa forma, a utilização de imagens isoladas com o *Detectron2* não foi a melhor alternativa para esse projeto, tendo em vista outros métodos que conseguiram obter keypoints de forma mais precisa, como o caso do artigo de Bridgeman, et al. (2019).

Já na tarefa de classificação de times, ao comparar os resultados do modelo de *clustering* utilizado com as classificações do *dataset*, obteve-se que o modelo cometeu erro de classificação em 19,5% dos jogadores corretamente identificados pelo modelo de detecção. Desses 19,5%, 8,7% são jogadores classificados erroneamente como não-pertencentes aos times 1 e 2, ou seja, são erros advindos da clusterização com o *DBSCAN*. 8,2% são erros de classificação entre time 1 ou time 2, ou seja, erros advindos da clusterização com o modelo de mistura gaussiana. Ainda há 0,1% de goleiros que não foram classificados como tal e 2,5% que foram classificados como jogadores, porém não estão no *dataset*. Dentro dessa última categoria de erro estão árbitros ou outras pessoas externas identificadas como jogadores do time 1 ou do time 2, mas também há casos de erros no *dataset*, ou seja, jogadores cujos keypoints não estão registrados. Dentre as demais causas de erro de classificação de time encontradas também estão as sombras, que modificam a cor da camisa do jogador, e imprecisões na posição das keypoints, que levam a erros na identificação da região da camiseta do jogador.

Por fim, ao combinar o erro de detecção dos jogadores com o erro de classificação de times (porém desconsiderando os 2,5% advindos do erro de classificação de pessoas fora do *dataset*), tem-se, ao total, que os jogadores do *dataset* foram detectados e tiveram o time identificado corretamente em 78,9%. Considerando a precisão completa das demais etapas da classificação de impedimento, essa taxa de acerto já seria suficiente para se equiparar com árbitros humanos, que possuem taxa de acerto perto de 75%, como apontado por Helsen, et al. (2007).

#### *B. Identificação do Ponto de Fuga e Traçado das Linhas de Impedimento*

A implementação da Transformada de Hough nas imagens de partidas de futebol provou ser eficaz na detecção das linhas de demarcação, cruciais para identificar o ponto de fuga. O método mostrou robustez em diversas condições de iluminação e qualidade de imagem. Após a filtragem das linhas, utilizamos um processo de minimização da soma das distâncias quadráticas a todas as interseções das linhas para encontrar o ponto de convergência.

Aliado à identificação dos jogadores, seus times e keypoints, a determinação do ponto de fuga foi essencial para traçar linhas de impedimento com precisão, demonstrando alinhamento adequado com as situações reais de jogo. De posse desses dados, a determinação da linha de impedimento se seguia naturalmente e sem custo computacional adicional.

## V. CONCLUSÃO

A utilização de técnicas de visão computacional na detecção de impedimentos no futebol representa a possibilidade de um avanço significativo na análise e arbitragem do jogo. Essa abordagem inovadora oferece benefícios consideráveis em termos de precisão e imparcialidade, melhorando a qualidade das decisões tomadas durante as partidas.

Ficou claro que ao empregar as técnicas propostas, foi possível capturar e processar rapidamente as informações visuais do campo, permitindo uma análise quase em tempo real das posições dos jogadores entre si. Essa automação do processo reduz a margem de erro humana e contribui para uma tomada de decisão mais justa e precisa.

É também importante identificar os principais gargalos e futuros pontos de pesquisa: na parte de identificação de jogadores a principal dificuldade acontecia quando dois jogadores ou mais eram identificados como um só, situação relativamente corriqueira em disputas desse tipo; também é possível sofisticar o modelo considerando a construção da pose tridimensional dos jogadores, fato que simplificaria a aplicação das heurísticas de projeção das partes do corpo consideradas no item *III-D*; já a determinação do ponto de fuga poderia ser mais rápida, agilizando o processo de traçado das possíveis linhas de impedimento.

De maneira geral, os resultados foram satisfatórios e apontaram para a real possibilidade de se automatizar o processo de impedimento, ou pelo menos sua diligência. Com um pouco mais de esforço seria possível implementar um algoritmo de tempo real para encontrar possíveis momentos de impedimento e analisá-los com base nas técnicas aqui apresentadas.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Panse N., Mahabaleshwarkar A., A dataset and methodology for computer vision based offside detection in soccer. **Proceedings of the 3rd International Workshop on Multimedia Content Analysis in Sports**, 19-26, 2020. <https://doi.org/10.1145/3422844.3423055>
- [2] Bridgeman L., et al. Multi-person 3D Pose Estimation and Tracking in Sports. **International Workshop on Computer Vision in Sports**, 5., ju. 2019.
- [3] Helsen W., et al. Errors in judging "offside" in association football: Test of the optical error versus the perceptual flash-lag hypothesis, **Journal of Sports Sciences**, 24:5, 521-528, 2006. doi: 10.1080/02640410500298065