



UNIVERSIDADE  
ESTADUAL DE LONDRINA

---

MURILO JOSÉ DE OLIVEIRA BUENO

**DIMENSÃO FRACTAL MULTIESCALA APLICADA À  
ANÁLISE TÁTICA NO FUTEBOL:  
UMA PROPOSTA DE AVALIAÇÃO E DESCRIÇÃO DE FORMAS  
DE ORGANIZAÇÃO COLETIVA DE ATLETAS EM CAMPO**

---

Londrina  
2022

MURILO JOSÉ DE OLIVEIRA BUENO

**DIMENSÃO FRACTAL MULTIESCALA APLICADA À  
ANÁLISE TÁTICA NO FUTEBOL:  
UMA PROPOSTA DE AVALIAÇÃO E DESCRIÇÃO DE FORMAS  
DE ORGANIZAÇÃO COLETIVA DE ATLETAS EM CAMPO**

Tese de Doutorado apresentada junto ao Programa de Pós-Graduação Stricto Sensu em Educação Física associado UEL/UEM, da Universidade Estadual de Londrina, como requisito para obtenção do título de Doutor em Educação Física.

Orientador: Prof. Dr. Felipe Arruda Moura

Londrina,  
2022

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

Bueno, Murilo José de Oliveira.

Dimensão fractal multiescala aplicada à análise tática no futebol: uma proposta de avaliação e descrição de formas de organização coletiva de atletas em campo / Murilo José de Oliveira Bueno. - Londrina, 2022.  
87 f.

Orientador: Prof. Dr. Felipe Arruda Moura.

Tese (Doutorado em Educação Física) - Universidade Estadual de Londrina, Centro de Educação Física e Esportes, Programa de Pós-Graduação em Educação Física, 2022.  
Inclui bibliografia.

1. Esportes coletivos - Tese. 2. análise de partidas - Tese. 3. dimensão fractal - Tese. 4. variabilidade - Tese. I. Moura, Prof. Dr. Felipe Arruda. II. Universidade Estadual de Londrina. Centro de Educação Física e Esportes. Programa de Pós-Graduação em Educação Física. III. Título.

CDU 796

MURILO JOSÉ DE OLIVEIRA BUENO

**DIMENSÃO FRACTAL MULTIESCALA APLICADA À  
ANÁLISE TÁTICA NO FUTEBOL:  
UMA PROPOSTA DE AVALIAÇÃO E DESCRIÇÃO DE FORMAS  
DE ORGANIZAÇÃO COLETIVA DE ATLETAS EM CAMPO**

Tese de Doutorado apresentada junto ao Programa de Pós-Graduação Stricto Sensu em Educação Física associado UEL/UEM, da Universidade Estadual de Londrina, como requisito para obtenção do título de Doutor em Educação Física.

**BANCA EXAMINADORA**

---

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

---

Prof. Dr. Allan da Silva Pinto  
Laboratório Nacional de Luz Sincrotron

---

Prof. Dr. Paulo Roberto Pereira Santiago  
Universidade de São Paulo - USP

---

Prof. Dr. Rodrigo Leal de Queiros Thomaz Aquino  
Universidade Federal do Espírito Santo - UFES

---

Prof. Dr. André Gustavo Pereira de Andrade  
Universidade Federal de Minas Gerais - UFMG

Londrina, 29 de Abril de 2022

## DEDICATÓRIA

Dedico este trabalho a uma pessoa que é/foi grande e maravilhoso pai, filho, irmão, primo, tio, avô, que tinha vários nomes (João Marcos Bueno, João Marcos, simplesmente “Joãozinho”, “Jão”; “tio doido”, entre outros...) e sempre era chamado assim com carinho. Para mim era meu pai, um excelente pai, amigo, parceiro de cervejas, de conversas sérias ou engraçadas. Sem meu pai eu não chegaria ao que sou hoje, não somente por apoio financeiro, mas um apoio paterno, de exemplo de pessoa que foi para mim e para minha irmã.

Meu pai, como eu sinto a falta de você, como eu gostaria que estivesse aqui em pessoa para ver o final deste trabalho, que sempre foi dedicado a você. Mas sei que onde estiver, e acredito na vida eterna, está contente por isso.

Eu te amo, “Jão”!

## AGRADECIMENTOS

Quando aqui eu apresento os agradecimentos, não busco hierarquizar ou apresentar uma ordem de importância, mas coloco em exceção abaixo:

Agradeço a Deus pela minha vida, em tudo o que proporciona para comigo, e assim busco aceitar a providência divina.

A partir deste ponto, todos em suma têm sempre uma importância em minha vida e para a conclusão deste trabalho.

Agradeço ao meu orientador pela amizade conquistada ao longo de vários anos. Pelo exemplo de trabalho em que realiza e transmite aos seus alunos. Pela paciência que teve para comigo durante todos os anos em que estive sob sua orientação. Também não posso deixar de agradecer em momentos pessoais que estive comigo, em apoio nos momentos de angústias e tristezas, e momentos de alegrias e comemorações. O meu muito obrigado por essa grande amizade!

Ao meu pai João Marcos Bueno (*in memoriam*) pelo que você representa em minha vida. Um excelente pai, amigo, companheiro, que em tudo, sempre pensou antes na família, principalmente em mim e minha irmã, como também seu neto, sem antes pensar em você. Foi e sempre será meu exemplo de pessoa quem eu quero ser. Eu te amo!

Quero agradecer à minha mãe Ana Maria de Oliveira pelos carinhos, abraços, acolhimentos, principalmente quando era mais novo. Cuidou de mim e de minha irmã e hoje nos tornamos bons adultos graças aos seus cuidados de mãe, amor de mãe.

À minha irmã Rafaela F. de Oliveira Bueno e a meu cunhado Rafael H. Muniz Cassetari, agradeço pelos seus carinhos, paciência comigo, seus abraços e amizade. Muito obrigado por tudo, amo vocês!

Agradeço a uma pessoa que hoje é pequena, tem 10 anos, não pode entender o significado de concluir esta tese, mas que muitas vezes, quando estava a escrever/estudar pensava nele em como devo ser um exemplo de aluno, professor, irmão, filho, mas principalmente, como eu devo ser como tio e amigo para com ele. Miguel Bueno Cassetari, o tio te ama!

Agradeço a todos os membros de minha Família, tios, primos, avós (*in memorian*) por todo o apoio e carinho em minha vida.

Quero agradecer a todos os membros do Laboratório de Biomecânica Aplicada – LBA pela amizade, paciência e grandes debates e contribuições para esta tese. Sem vocês não seria possível concluir. Muito obrigado!

Agradeço à Maisa Silva pela paciência em me socorrer quanto às muitas dúvidas quando o assunto era computação. Meus agradecimentos pela contribuição na tese e amizade.

Agradeço a todos os membros da banca pelas valiosas contribuições e sugestões para o término desta tese.

Agradeço em especial a dois professores que me auxiliaram durante estes anos que permearam o doutorado, e antes até. Agradeço ao Prof. Dr. Sergio Augusto Cunha e o Prof. Dr. Ricardo da Silva Torres, por grandes contribuições não somente para a tese, mas para a minha formação.

Agradecer à CAPES pelo auxílio financeiro para que esta tese pudesse ser concluída.

*“Bem-aventurados os pobres no espírito, pois deles é o Reino dos Céus.  
Bem-aventurados os que choram, pois eles serão consolados.  
Bem-aventurados os mansos, pois eles herdarão a terra.  
Bem-aventurados os que têm fome e sede da justiça, pois eles serão saciados.  
Bem-aventurados os misericordiosos, pois eles alcançarão misericórdia.  
Bem-aventurados os puros de coração, pois eles verão a Deus.  
Bem-aventurados os que promovem a paz, pois eles serão chamados filhos de Deus.  
Bem-aventurados os perseguidos por causa da justiça, pois deles é o reino dos Céus.  
Bem-aventurados sois vós, quando vos injuriarem e perseguirem e, mentindo, disserem todo  
mal contra vós por causa de mim.”*

*(Mateus, Cap. 5, 3 – 11).*

BUENO, Murilo José de Oliveira. **Dimensão fractal multiescala aplicada à análise tática no futebol**: uma proposta de avaliação e descrição de formas de organização coletiva de atletas em campo. 2022. 87 f. Tese (Doutorado em Educação Física) – Universidade Estadual de Londrina, Londrina, 2022.

## RESUMO

O objetivo desta tese foi desenvolver uma ferramenta por abordagem de Dimensão Fractal multiescala (DFM) para identificar e analisar as formas de organização de equipes de futebol durante partidas oficiais. Para tanto, dois estudos foram empregados: o primeiro abordou a análise da eficácia de diferentes descritores de forma (área/perímetro, área fractal, circularidade, máximo fractal, retangularidade, curva fractal multiescala – CFM), e a concatenação de todos os descritores de formas (exceto CFM), denominado *Alldescriptors* (AllID), e assim, avaliados e aplicados para polígonos correspondentes a formas, representadas pelo envoltório convexo obtido pelas coordenadas 2D dos jogadores. Para a avaliação dos descritores, um sistema de recuperação de imagens baseado em conteúdo (SRIBC) foi desenvolvido e medidas de eficácia foram utilizadas para avaliar os descritores de formas ( $P@n$  and  $R@n$ ). A CFM foi o mais eficaz descritor de formas em relação aos outros avaliados. Ao elucidar qual o melhor descritor de formas para a análise da organização de equipes (CFM), foi realizado o segundo estudo com demonstração de possíveis aplicações para descrição e análise de formas de organização de equipes de futebol. O descritor de formas Dimensão Fractal Multiescala (DFM) foi calculado e foi realizada a discriminação das formas de organização das equipes entre os períodos e por meio da frequência relativa de formas para cada *cluster* quando realizaram finalizações ou desarmes. Um Índice de Variabilidade (IV) de formas foi calculado pelo método *Sample Entropy* (SampEn) para as sequências ofensivas (SO) em situações de defesa e ataque, bem como a interação entre as formas pela Distância Euclidiana entre as Curvas Fractais Multiescala (DECFM). O Máximo valor de Distância Euclidiana (MvDE) entre as equipes também foi extraído. Com esta tese, foi possível elucidar que a DFM foi o melhor descritor de formas para descrever a organização de equipes de futebol, e assim, foi possível identificar as formas de organização em partidas oficiais de futebol. Ao aplicar o descritor DFM, pôde-se concluir que o sucesso ou o fracasso em sequências ofensivas em situação de defesa e ataque não advém da variabilidade de formas, porém, uma maior variabilidade de formas relacionadas à interação entre as equipes para as SO que terminaram em desarmes pode indicar uma possibilidade de sucesso durante as SO e uma imprevisibilidade na mudança de forma em sucesso na finalização.

**Palavras-chave:** esportes coletivos; análise de partidas; tática; dimensão fractal; recuperação de imagens baseado em conteúdo; variabilidade.

BUENO, Murilo José de Oliveira. **Multiscale fractal dimension applied to tactical analysis in football**: a proposal for the evaluation and description of team organization shapes on the pitch. 2022. 87 p. Thesis (Doctoral Degree in Physical Education) – Universidade Estadual de Londrina, Londrina, 2022.

## ABSTRACT

The objective of this thesis was to develop a tool using a multiscale Fractal Dimension (MFD) approach to characterize the shapes of tactical organization of football teams during official matches. Two studies were developed: the first addressed the analysis of the effectiveness of different shape descriptors (area/perimeter, fractal area, circularity, maximum fractal, rectangularity, multiscale fractal curve – MFC), and the concatenation of all shape descriptors (except MFC), denominated Alldescriptors (AlID), and thus, evaluated and applied to polygons corresponding to shapes, represented by the convex hull obtained by the 2D coordinates of the players. For the evaluation of the descriptors, a content-based image retrieval system (SRIBC) was developed and efficacy measures were used to evaluate the shape descriptors (P@n and R@n). The MFC was the most effective shape descriptor in relation to the others evaluated. When elucidating the best descriptor of shapes for the analysis of team organization (MFC), the second study was carried out with a demonstration of possible applications for the description and analysis of shapes organization of football teams. The Multiscale Fractal Dimension (MFD) shape descriptor was calculated and the shapes of organization of the teams were discriminated between the periods and through the relative frequency of shapes for each cluster when they performed shots on goal or tackles. A Variability Index (IV) of shapes was calculated by the Sample Entropy method (SampEn) for the offensive sequences (SO) in defense and attack situations, as well as the interaction between the shapes by the Euclidean Distance between the Multiscale Fractal Curves (DECFM). The Maximum Euclidean Distance (MvDE) value between teams during SO was also extracted. With this thesis, it was possible to elucidate that the MFD was the best shape descriptor to describe the organization of football teams, and thus, it was possible to identify the shapes of organization in official football matches. Applying the MFD descriptor, it was possible to conclude that the success or failure in offensive sequences in defense and attack situations do not come from the variability of shapes, however, a greater variability of shapes related to the interaction between the teams for the SO that ended on tackles can indicate a chance of success during SO and the unpredictability of shape change in successful on shots on goal.

**Keywords:** team sports; match analysis; tactics; fractal dimension; content-based image retrieval; variability.

## LISTA DE FIGURAS

- Figura 1** – Posicionamento das câmeras para posterior rastreamento dos jogadores. (Fonte: o próprio autor). ..... 28
- Figura 2** – Interface para realização da segmentação dos jogos e identificação dos jogadores. (Fonte: o próprio autor). ..... 29
- Figura 3** – Ilustração dos pontos de obtenção das métricas reais no campo (C1, C2,...Cn). (Fonte: o próprio autor). ..... 30
- Figura 4** – Sequências de imagens no jogo (A); Obtenção dos blobs correspondente aos jogadores (B); Representação do grafo construído a partir dos blobs (C). (Fonte: o próprio autor). ..... 31
- Figura 5** – Identificação da posição de cada jogador em campo ao longo de uma sequência de imagens. (Fonte: o próprio autor). ..... 31
- Figura 6** – Interface para o registro das ações técnicas realizadas pelos jogadores. (Fonte: o próprio autor). ..... 32
- Figura 7** – Etapas para a construção das imagens de polígonos convexos que representam a organização das equipes em campo: (a) posição dos jogadores de cada equipe. (b) Aplicação do método *quickhull* para a construção do envoltório convexo. (c) Polígono do envoltório convexo na respectiva posição em campo. Imagens dos formatos analisados para cada equipe, (d) time 1 e (e) time 2. (Fonte: o próprio autor). ..... 33
- Figura 8** – (a) Imagem de um polígono; (b) Função logarítmica da área; (c) Dimensão Fractal Multiescalar da imagem analisada (Máximo Valor Fractal (MF); Curva Fractal Multiescala (CFM), área fractal (AF)). (Fonte: o próprio autor). ..... 35
- Figura 9** – Exemplo da interface para a avaliação da relevância dos polígonos em relação ao *Query* (imagem consulta). Abaixo os 6 polígonos para que o usuário possa avaliar sua relevância. Em azul um exemplo de polígonos possivelmente relevantes identificados pelo usuário. (Fonte: o próprio autor). ..... 45
- Figura 10** – Representação dos resultados de desempenho de busca em relação a curva Precision x Recall para todos os descritores avaliados. (Fonte: o próprio autor). ..... 48

- Figura 11** – Resultado do agrupamento da CFM para as equipes 1 e 2: (A) e (B), 1° e 2° período da equipe 1, respectivamente; (C) e (D), 1° e 2° período da equipe 2, respectivamente. Resultados também das formas encontradas para cada cluster analisado, quando ocorrem ao longo da partida (gráfico de barras) e percentis de formas em cada cluster (gráfico de pizza). (Fonte: o próprio autor). ..... 51
- Figura 12** – Esquema para a obtenção da série temporal correspondente a DECFM para as SO que resultaram em desarme ou finalização. (a), obtenção dos polígonos convexos; (b), obtenção da CFM e cálculo da DE; (c), exemplo de série temporal da DECFM e o MvDE para uma SO. (Fonte: o próprio autor)..... 63
- Figura 13** – Exemplos de formas de organização, percentual (*pie graphics*), e momento de ocorrências de formas (*bar graphics*) em uma partida. (A), equipe anfitriã, 1° período; (B), equipe anfitriã, 2° período; (C), equipe visitante, 1° período; e (D) equipe visitante, 2° período. (Fonte: o próprio autor)..... 65
- Figura 14** – Resultado de Frequência Relativa (FR) de formas de organização, para cada cluster, que as equipes realizam em momentos específicos de situações de ataque e defesa. (A), FR em situações de defesa, 1° período; (B), FR em situações de defesa, 2° período; (C), FR em situações de ataque, 1° período; e (D), FR em situações de ataque, 2° período. (Fonte: o próprio autor)..... 66
- Figura 15** – Exemplo de séries temporais da distância Euclidiana das CFM entre as equipes durante uma SO que terminou em desarme (a), e uma SO que terminou em finalização (b). (Fonte: o próprio autor). ..... 68

## LISTA DE TABELAS

- Tabela 1** – Resultados das análises de concordância (Coeficiente de Kappa) para identificação de relevância dos formatos para cada descritor entre os 3 diferentes avaliadores (A, B, e C).....47
- Tabela 2** – Resultados do percentual de concordância entre imagens consideradas não relevantes (NR - NR) e imagens consideradas relevantes (R - R) entre 3 diferentes avaliadores. ....47
- Tabela 3** – Valores de média  $\pm$  desvio padrão para precision (P@5 e P@10) e recall (R@5 e R@10) para avaliar a eficácia de diferentes descritores. ....49
- Tabela 4** – Resultados de mediana (intervalo interquartil) para os valores de IV de formas (SampEn) em sequências ofensivas (SO) para situações de defesa (n = 1193) e situações de ataque (n = 256)..... 67
- Tabela 5** – Resultados de mediana (intervalo interquartil) dos valores de IV (SampEn) e o máximo valor extraído para os dados de Distância Euclidiana entre as Curvas Fractais Multiescala (DECFM) das equipes para as SO que terminaram em finalizações e desarmes. .... 68

## LISTA DE QUADROS

<b>Quadro 1</b> – Cálculo de descritores de formas Retangularidade, Circularidade e Área/Perímetro para exemplos de formas genéricas.....	36
---	----

## LISTA DE ABREVIATURAS E SIGLAS

AF	Área Fractal
AlID	<i>All descriptors</i>
CFM	Curva Fractal Multiescala
DE	Distância Euclidiana
DECFM	Distância Euclidiana entre a Curva Fractal Multiescala
DF	Dimensão Fractal
DFM	Dimensão Fractal Multiescala
FR	Frequência Relativa
IV	Índice de Variabilidade
MF	Máximo valor Fractal
MvDE	Máximo valor de Distância Euclidiana
SO	Sequências Ofensivas
VC	Vetor de Características

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO GERAL</b> .....	20
<b>2</b>	<b>ORGANIZAÇÃO DA TESE</b> .....	24
<b>3</b>	<b>OBJETIVOS</b> .....	26
3.1	OBJETIVO GERAL .....	26
3.2	OBJETIVOS ESPECÍFICOS .....	26
<b>4</b>	<b>PROCEDIMENTOS METODOLÓGICOS</b> .....	27
4.1	COLETA DE DADOS: OBTENÇÃO DA POSIÇÃO 2D DOS JOGADORES E AÇÕES TÉCNICAS DURANTE AS PARTIDAS OFICIAIS .....	27
4.2	OBTENÇÃO DAS FORMAS DE ORGANIZAÇÃO DAS EQUIPES DURANTE AS PARTIDAS OFICIAIS .....	32
4.3	DESCRITORES DE FORMAS .....	33
<b>5</b>	<b>ESTUDO UM: DIMENSÃO FRACTAL MULTIESCALA APLICADA ÀS ANÁLISES TÁTICAS NO FUTEBOL: UMA NOVA ABORDAGEM PARA AVALIAR AS FORMAS DE ORGANIZAÇÃO DE EQUIPES EM CAMPO</b> .....	38
5.1	INTRODUÇÃO .....	39
5.2	MATERIAIS E MÉTODOS .....	40
5.2.1	Coleta De Dados .....	40
5.2.2	Elaboração Dos Polígonos Que Representam A Organização Tática De Equipes De Futebol. ....	41
5.2.3	Descritores De Formas .....	41
5.2.4	Sistema De Recuperação De Imagens Baseado Em Conteúdo (SRIBC): Etapas Para A Implementação .....	43
5.2.5	Avaliação Da Eficácia De Descritores De Formas. ....	44
5.2.6	Análises Estatísticas. ....	46
5.2.7	Resultados .....	47
5.2.8	Exemplo De Aplicação Utilizando Descritores De Formas Para Análises De Partida No Futebol .....	49

5.3	DISCUSSÃO.....	52
5.4	CONCLUSÃO.....	55
<b>6</b>	<b>ESTUDO DOIS: QUAIS SÃO AS FORMAS DE ORGANIZAÇÃO QUE EQUIPES DE FUTEBOL ADOTAM E SUA RELAÇÃO COM EVENTOS DO JOGO? IDENTIFICAÇÃO E ANÁLISE EM PARTIDAS OFICIAIS .....</b>	<b>56</b>
6.1	INTRODUÇÃO.....	57
6.2	MATERIAIS E MÉTODOS .....	60
6.2.1	Participantes .....	60
6.2.2	Coleta De Dados.....	60
6.2.3	Análise Dos Dados .....	62
6.2.4	Análises Estatísticas .....	63
6.3	RESULTADOS.....	64
6.4	DISCUSSÃO.....	68
6.5	CONCLUSÃO .....	71
<b>7</b>	<b>CONSIDERAÇÕES FINAIS.....</b>	<b>72</b>
<b>8</b>	<b>REFERÊNCIAS .....</b>	<b>74</b>
	<b>APÊNDICES.....</b>	<b>79</b>
	<b>ANEXOS .....</b>	<b>83</b>

## 1 Introdução geral

Pesquisas sobre o universo de esportes coletivos, em particular o futebol, buscam obter e compreender uma gama de variáveis capazes de aperfeiçoar o desempenho do atleta ou equipe. Pesquisadores, principalmente das ciências dos esportes, buscam esse aperfeiçoamento metodológico para, no caso do futebol, poder compreender todas (ou quase todas) as ações de um determinado jogador durante o jogo e fornecer informações para que técnicos/professores possam elaborar estratégias para melhora do desempenho do atleta ou equipe (Bueno et al., 2018). Variáveis como distância percorrida total dos jogadores, distâncias percorridas por faixas de velocidades (Barros et al., 2007; Di Salvo et al., 2007; Di Salvo et al., 2010; Dellal et al., 2011) podem ser analisadas durante as partidas, bem como a distribuição tática de jogadores da mesma equipe e adversária (Frencken et al., 2011; Moura et al., 2012).

Com o avanço no desenvolvimento de ferramentas computacionais aplicadas às análises de dados posicionais, em proporção ao *big data*, torna-se possível estabelecer como o comportamento tático de equipes em jogo pode ser representado e analisado durante as partidas (Rein e Memmert, 2016; Goes et al., 2020). Vários métodos computacionais têm sido desenvolvidos e aplicados para compreender o comportamento tático em esportes coletivos relacionados a parâmetros espaciais (Rico-González et al., 2020(a); Rico-González et al., 2020(b); Rico-González et al., 2021). Estas variáveis são apresentadas na literatura como componentes relevantes para as análises na organização de equipes de futebol em campo, como centroide da equipe, medidas de espalhamento, *stretch index*, largura, comprimento e área de ocupação que as equipes ocupam em campo (Bartlett et al., 2012; Moura et al., 2012; Duarte R. et al., 2013; Clemente M. F. et al., 2017), bem como as regiões de domínio de cada jogador em campo (Fonseca et al., 2013; Caetano et al., 2021).

Pesquisas realizadas no domínio das ciências dos esportes e ciências da computação muitas vezes são abordadas sob diferentes paradigmas. Goes et al. (2020) elaboraram, por meio de uma revisão sistemática, uma estrutura conceitual que mostra a contribuição de ambas ciências para conhecimentos sobre o comportamento tático no futebol. As ciências dos esportes se preocupam principalmente com questões práticas (intervenções representam 63% das pesquisas), seguidas por problemas metodológicos (20%), e relacionais (17%), por exemplo, magnitudes de área de ocupação e espalhamento relacionadas ao sucesso de ações

determinantes no futebol, como finalização ou desarme (Moura et al., 2012). Quando observados estudos que utilizaram e analisaram as variáveis táticas aplicadas para compreender a organização das equipes, a contribuição das ciências dos esportes foi considerada maior para as variáveis de distribuição dos jogadores, como área de ocupação e espalhamento (Goes et al., 2020), o que mostra uma aplicação maior para solucionar problemas relacionados à prática da modalidade e menor ênfase para desenvolvimento metodológico.

Na literatura, pode-se observar que a área de ocupação é uma das possibilidades para representar como as equipes se distribuem durante os jogos (Frencken e Lemmink, 2009; Frencken et al., 2011; Moura et al., 2012), dentre as variáveis citadas. A área de ocupação é normalmente definida como a área do menor polígono convexo formado pela posição dos jogadores, e representa o espaço ocupado por uma equipe em campo (Frencken e Lemmink, 2009; Moura et al., 2012). Esses valores em situações defensivas e ofensivas, para ações como finalizações ou desarmes, são distintas quando ocorrem o sucesso ou o fracasso entre estas ações (Moura et al., 2012). Essas diferenças podem estar atreladas aos objetivos da equipe. Quando defendem, buscam compactar, evitar uma finalização e recuperar a posse de bola. Quando atacam, os jogadores se organizam para buscar oportunidades para chegar a sua meta, e realizar um chute a gol (Mitchell, 1996).

Apesar de a literatura, principalmente a ciência dos esportes, considerar os valores de magnitude da área de ocupação bons para a caracterização da dinâmica de equipes em esportes coletivos durante uma partida, argumenta-se que diferentes formas<sup>1</sup> de organizações de jogadores em campo podem fornecer áreas de ocupação iguais ou o oposto; diferentes áreas de ocupação podem apresentar formas semelhantes. Neste sentido, uma análise mais detalhada sobre a forma do polígono em que uma equipe apresenta em campo associada à área de ocupação e sua complexidade, pode fornecer informações adicionais sobre características estruturais que as equipes apresentam durante uma partida. A forma estrutural é considerada uma importante característica útil para descrever e distinguir determinados objetos (Locaric, 1998), sendo possível tal aplicação para análises de organização no futebol.

Alguns descritores de formas são apresentados na literatura, basicamente classificados em duas categorias: análises baseadas em contorno ou baseadas em região (Zhang e Lu,

---

<sup>1</sup> Neste documento, “forma” ou “formato” serão tratados como sinônimos, que correspondem a uma parte da morfologia de figuras geométricas correspondentes à organização de equipes de futebol em campo.

2004). Conforme seus nomes correspondentes, o primeiro usa apenas informações do contorno do objeto, enquanto o último emprega todos os *pixels* de imagem para a obtenção de um vetor de característica.

Uma abordagem utilizada para estimar a complexidade de uma determinada forma consiste no uso de medidas de dimensão fractal. A geometria fractal (Mandelbrot, 1982) é um campo da matemática com objetivo de analisar a complexidade de uma forma. Esta teoria tem sido utilizada em processamento de imagens, especificamente para segmentação de texturas (Chaudhuri e Sarkar, 1995) e compressão de imagens (Fisher, 1995). Uma de suas definições fundamentais é a dimensão fractal, uma medida de complexidade que generaliza o conceito de dimensão topológica. A Dimensão Fractal Multiescala (DFM), por sua vez, é uma extensão da dimensão fractal *Minkowski-Bouligand*. Este utiliza a dilatação morfológica euclidiana para representação multiescala (Torres, Falcão e Costa, 2004).

Portanto, além de tratar as magnitudes de área de ocupação e suas variações ao longo da partida, a estrutura (forma) de organização pode ser imprescindível para compreensão de como atletas e equipes se organizam frente aos objetivos do jogo. Equipes de futebol são, portanto, dois sistemas complexos com auto-organização, que interagem, se adaptam e variam ao longo do tempo, o que também representa um fator de desempenho (Frencken e Lemmink, 2009; Moura et al., 2016). Assim, uma vez que as equipes se adaptam e variam seu comportamento ao longo da partida para atingirem seus objetivos, pode-se questionar como a variabilidade, não de magnitude mas da forma de organização, pode explicar o sucesso ou fracasso durante jogadas de ataque e defesa.

A variabilidade é estudada em diversas áreas de conhecimento, principalmente em pesquisas relacionadas a sistemas dinâmicos, biomecânica, aprendizagem e controle motor (Bartlett, Wheat e Robins, 2007). Os vários graus de liberdade envolvidos na coordenação e o controle do movimento são uma fonte potencial de variabilidade. Em biomecânica e controle motor, a variabilidade é tradicionalmente equiparada ao ruído, e considerado prejudicial ao desempenho de um sistema, e normalmente é eliminado dos dados como uma fonte de erro (Van Emmerik, Miller e Hamill, 2014). No entanto, as ciências biológicas e físicas enfatizam os aspectos benéficos e adaptativos no funcionamento do sistema. Na perspectiva tradicional, níveis superiores de habilidade, competência e saúde estão associados com diminuição da variabilidade. Porém, perspectivas atuais de sistemas dinâmicos apontam o oposto; a perda de variabilidade reflete uma diminuição da interação dos graus de liberdade, o que representa um

prejuízo no funcionamento do sistema (Van Emmerik, Miller e Hamill, 2014). Um estudo recente mostrou a variabilidade de posicionamento de atletas ao longo de uma competição renomada da Europa e mostrou que as equipes finalistas apresentam menor variabilidade de posição média dos jogadores ao longo do campeonato (Moura et al., 2015). No entanto, a posição média dos jogadores, em uma partida, ainda não reflete de fato a capacidade da equipe em se organizar ao longo da partida.

Assim, esta pesquisa baseou-se na necessidade em responder algumas lacunas correspondentes à organização de equipes de futebol em partidas oficiais. Conforme será apresentado no decorrer desta tese, existem vários métodos e variáveis que podem quantificar a magnitude de ocupação e organização das equipes de futebol em campo durante partidas oficiais. Porém, observa-se uma lacuna na literatura na qual a forma e complexidade de organização que as equipes atuam durante as partidas não são exploradas e analisadas.

Até onde se sabe, não foi encontrado nenhum estudo ou método que descreva e analise a forma de organização de equipes de futebol e que utilize descritores baseados em forma. Assim, ao analisar a forma com que as equipes se organizam em campo, é possível descrever os padrões de organização que as mesmas ocupam em campo, compreender o quão complexo são as organizações das equipes em condições de ataque e defesa, e se esta complexidade pode apresentar diferentes modificações ou variabilidade que levam as equipes ao sucesso ou fracasso em ações determinantes da partida como finalizações ou desarmes. Devido à escassez de ferramentas já citada acima, faz-se necessária a implementação, avaliação e validação de métodos de descritores de forma, que possam melhor descrever esses polígonos que correspondem à organização de equipes de futebol em campo.

## 2 Organização da tese

As etapas pertinentes da presente tese foram organizadas e apresentadas em distintos capítulos, a saber:

– ***Procedimentos metodológicos:*** É destinado à descrição detalhada da obtenção da posição 2D dos jogadores durante as partidas oficiais de futebol, bem como o registro das ações técnicas e posse de bola dos jogadores (quando pertinente). No capítulo, são apresentadas as etapas para a obtenção das imagens do polígono convexo que representa a organização das equipes analisadas, bem como os descritores de formas (DF) avaliados nesta tese: área/perímetro, circularidade, retangularidade, curva fractal multiescala (CFM), área fractal multiescala, máximo valor fractal e por último a concatenação de todos os descritores avaliados, denominado *Alldescriptors* (AIID), com ressalva à CFM. Estes descritores são melhores elucidados no capítulo correspondente.

– ***Estudo um: Dimensão fractal multiescala aplicada às análises táticas no futebol: uma nova abordagem para avaliar as formas de organização de equipes em campo:*** Apresenta as etapas destinadas à criação de um sistema de busca de imagens baseado em conteúdo (SIBC) para a avaliação de eficácia dos descritores de formas avaliados nesta tese. Posteriormente, são apresentados os resultados de eficácia dos DF avaliados aplicados à busca de imagens de polígonos formados pelo envoltório convexo que representam a organização no futebol, e a aplicação do(s) melhor(es) descritor(es) avaliado(s) para identificar quais formas as equipes realizam durante uma partida de futebol. Neste, um método de agrupamento, *k-means*, é aplicado para identificar as formas como uma possibilidade de aplicação. Neste estudo, as questões norteadoras do problema de pesquisa foram:

– Dentre os descritores de formas mais utilizados na literatura, qual(is) é(são) o(s) melhor(es) descritor(es) para descrever a forma de organização de equipes ao longo de uma partida?

– ***Estudo dois: Quais são as formas de organização que equipes de futebol adotam e sua relação com eventos do jogo? Identificação e análise em partidas oficiais:*** Este capítulo apresentou uma aplicação da DFM para a descrição das formas de organização tática em partidas oficiais de futebol e também em janelas específicas de tempo que correspondem às sequências ofensivas (SO) que terminam em sucessos ou fracassos para situações de ataque e defesa. São apresentados os resultados de Índice de Variabilidade (IV) de formas por meio do

método *Sample Entropy* (SampEn) para as análises de SO em situação de ataque e defesa. Neste estudo, as questões norteadoras da pesquisa foram:

- Como equipes mudam suas formas de organização em relação ao adversário para atingir seus objetivos na partida?
- Qual é a relação entre a forma de organização e eventos determinantes da partida?
- Qual é o papel da variabilidade de forma em sequências ofensivas determinantes na partida?

Ao elucidar esses problemas de pesquisa, a proposta da tese foi desenvolver uma nova ferramenta para análise de formas de organização de equipes de futebol em partidas oficiais por meio de descritores de formas, e especificamente mostrar que as formas de organização das equipes são dinâmicas e apresentam diferentes escalas, o que faz necessária a utilização de descritores que considerem a variância de escala de formas estruturais que as equipes apresentam. Assim, esperou-se que a Dimensão Fractal Multiescala (DFM) poderia apresentar bons rendimentos para a descrição de formas de organização de equipes de futebol. A partir destes achados, podem-se apresentar novas aplicações para compreender a dinâmica de organização de acordo com as estruturas que as equipes apresentam, por exemplo, em situações de ataque e defesa, e como a variabilidade de formas pode ter relação com sucessos de ações determinantes no futebol.

### **3 Objetivos**

#### **3.1 Objetivo geral**

- Avaliar a abordagem de descritores de formas para caracterizar as formas de organização tática de equipes de futebol durante partidas oficiais.

#### **3.2 Objetivos Específicos**

Para que o objetivo geral da presente tese seja elucidado, os seguintes objetivos específicos foram determinados:

- Desenvolver e avaliar a eficácia de uma ferramenta capaz de identificar diferentes formas de organização de equipes de futebol por meio de descritores de formas;
- Avaliar diferentes descritores de formas aplicados às imagens de polígonos que representam a organização de equipes de futebol em campo e estabelecer qual(is) o(s) melhor(es) descritor(es) para descrever as formas de organização de equipes ao longo de uma partida;
- Identificar e analisar quais formas de organização que equipes de futebol realizam durante partidas oficiais por meio de descritores de formas, bem como em janelas específicas de tempo que correspondem às sequências ofensivas que terminam em sucessos ou fracassos para situações de ataque e defesa.

## 4 Procedimentos metodológicos

### 4.1 Coleta de dados: Obtenção da posição 2D dos jogadores e ações técnicas durante as partidas oficiais

Para este estudo, a amostra constituiu na obtenção de coordenadas 2D de jogadores de futebol profissional durante partidas oficiais de campeonatos brasileiro e estadual. Para tanto, um (1) jogo do campeonato paulista (Campeonato regional Série A-1) de 2004 foi coletado (n = 28 jogadores), nove jogos do campeonato brasileiro (Campeonato nacional Série A-1) foram coletados entre os anos de 2001/2003/2004/e2008 (n = 249 jogadores). Também, 4 jogos do campeonato paulista (Campeonato regional Série A-2) de 2014 (n = 101 jogadores) foram coletados. As descrições das partidas, locais, número de câmeras utilizadas para cada jogo e frequência de aquisição (7,5 ou 30 Hz) utilizadas para as coletas estão mais detalhadas no Apêndice I. Os jogos foram filmados por até 6 câmeras, posicionadas em locais altos dos estádios e permaneciam fixas do início ao fim das partidas. Cada câmera enquadrava aproximadamente um quarto do campo, assim, as imagens ficaram sobrepostas às outras (**Figura 1**). Após as filmagens, as imagens foram transferidas para computadores em *.avi* (*áudio vídeo interleaved*) e realizado o processo para o rastreamento automático dos jogadores no *software* DVide. O presente estudo foi aprovado pelo comitê de ética em pesquisa da Universidade Estadual de Londrina (parecer nº 3.047.461), conforme anexo I.

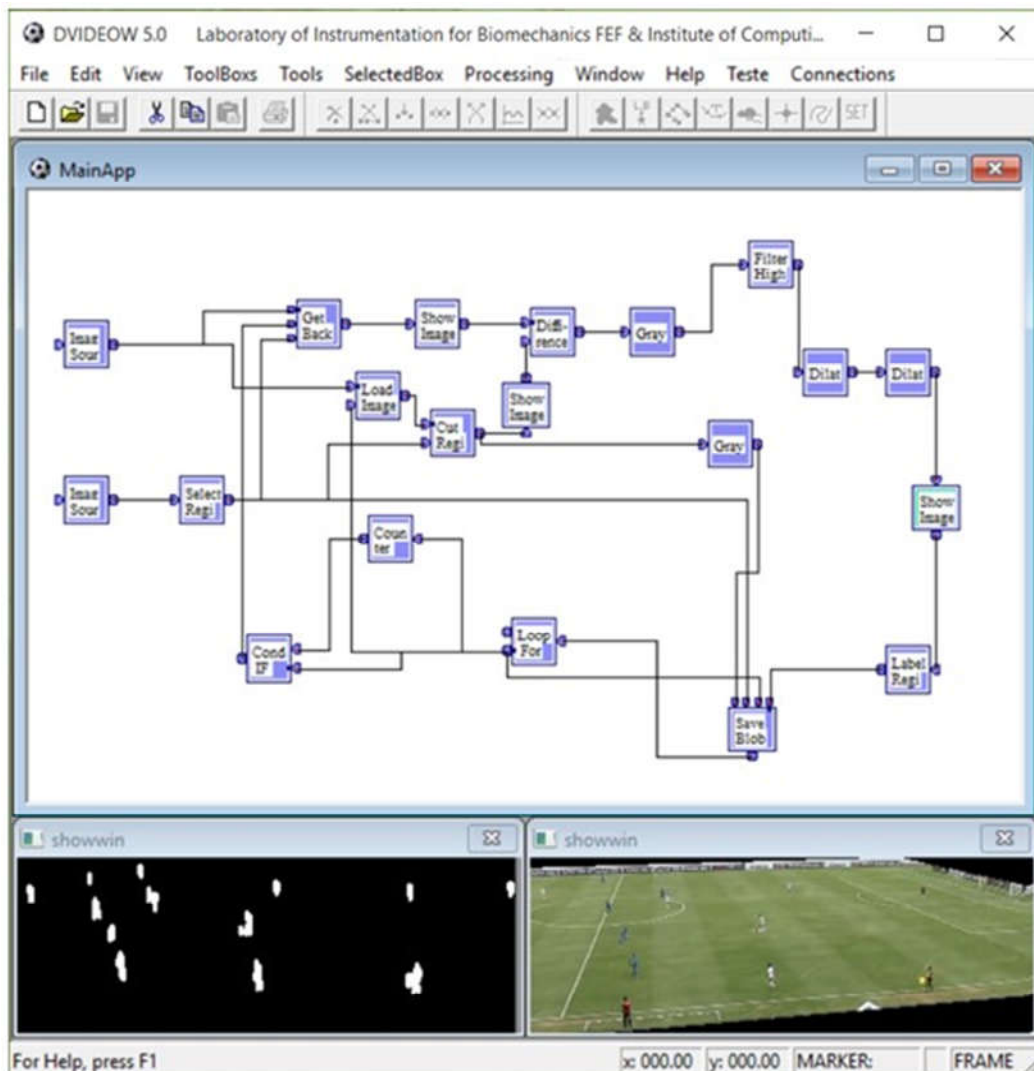
A partir das filmagens, foram obtidas as posições dos jogadores em campo por meio do *software* DVide (Figuerola, Leite e Barros, 2006a; 2006b), que utiliza ferramentas de visão computacional para a obtenção das coordenadas 2D dos jogadores. O *software* possui uma média de erro de 0,3 m para a determinação da posição do jogador e uma média de erro de 1,4% para a distância percorrida pelos jogadores (Figuerola, Leite e Barros, 2006b; Misuta, 2009). Esta medição pode ocorrer de forma manual ou automática. A medição manual ocorre quando o *software* não é capaz de identificar automaticamente o jogador de interesse ou quando identifica outro jogador erroneamente, o que ocorre em situações de oclusões ou quando há jogadores aglomerados. A realização da medição de forma automática envolve alguns procedimentos que identificam os jogadores em campo: a sincronização das câmeras, segmentação das imagens, calibração, construção do grafo e o rastreamento automático, conforme descrito abaixo:

- **Sincronização:** Foram utilizadas câmeras posicionadas em pontos mais altos do campo, de forma que as imagens de cada câmera fossem sobrepostas para enquadrar o campo todo (**Figura 1**). Estas imagens foram sincronizadas por meio de eventos em comum identificados entre as imagens sobrepostas (ex. um chute ou um passe que pudesse ser visualizadas por todas as câmeras).



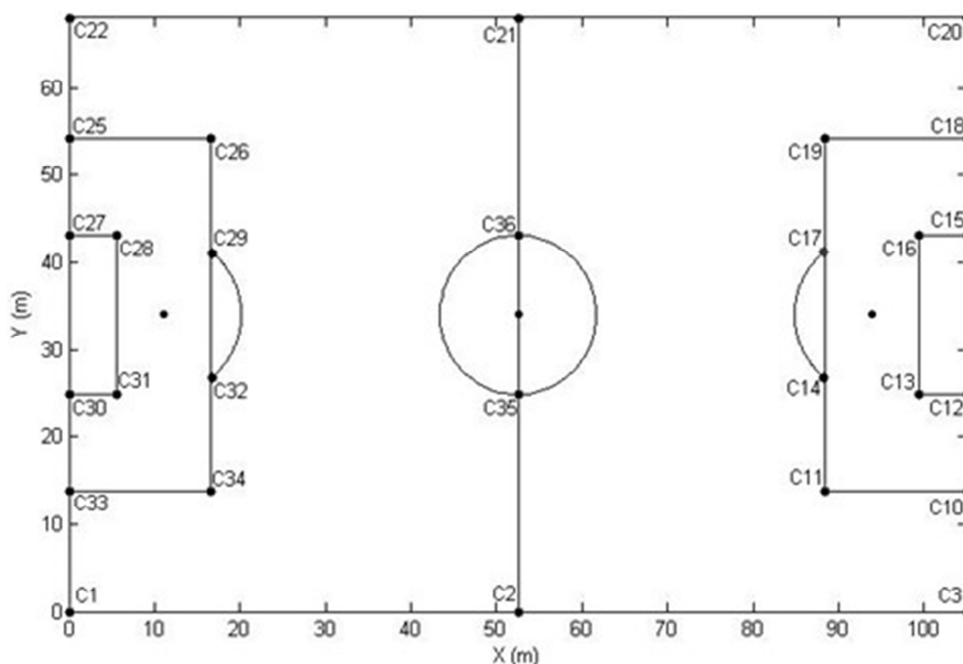
**Figura 1.** Posicionamento das câmeras para posterior rastreamento dos jogadores. (Fonte: o próprio autor).

- **Segmentação:** tem como objetivo detectar e separar os objetos de interesse (jogadores), a partir de uma sequência de imagens, dos outros elementos irrelevantes para a análise, como torcedores, placas de propaganda, quadra, entre outros. O *software* considera o relevo topográfico definido pelo jogador em determinada região do campo através de uma série de procedimentos computacionais, editados em sua interface específica para este fim. Após todo este processo, são obtidas imagens binarizadas e as regiões conexas encontradas são rotuladas e recebem o nome de *blobs*, conforme a **Figura 2**.



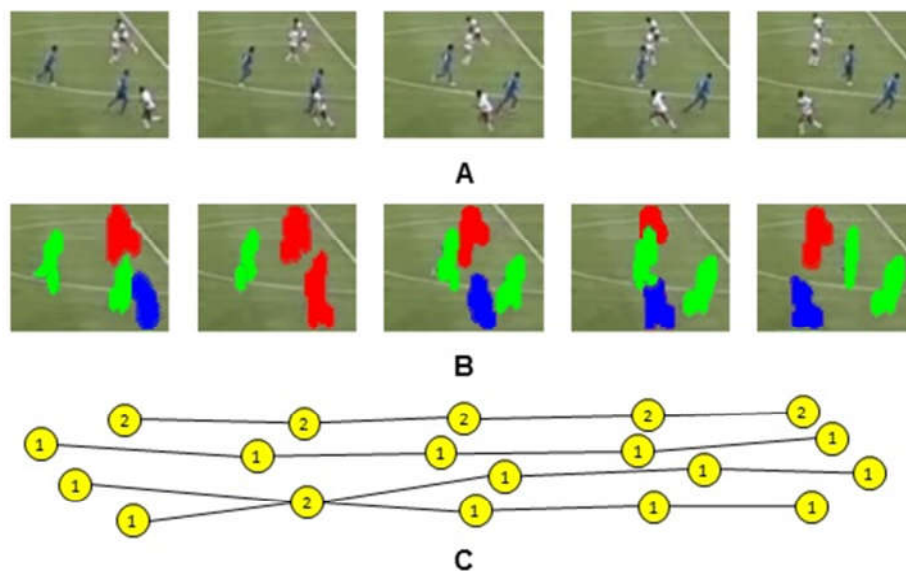
**Figura 2.** Interface para realização da segmentação dos jogos e identificação dos jogadores. (Fonte: o próprio autor).

- Calibração:** Antes de cada partida, foram obtidas as coordenadas de um conjunto de pontos específico em relação ao sistema de coordenadas associado ao campo. As projeções correspondentes desses mesmos pontos na imagem também foram determinadas com o *software* DVideo. Em seguida, os parâmetros homográficos da transformação objeto-imagem foram calculados com base na *Direct Linear Transformation* (DLT) e as coordenadas 2D dos jogadores em relação ao sistema de coordenadas do campo foram obtidas (**Figura 3**).

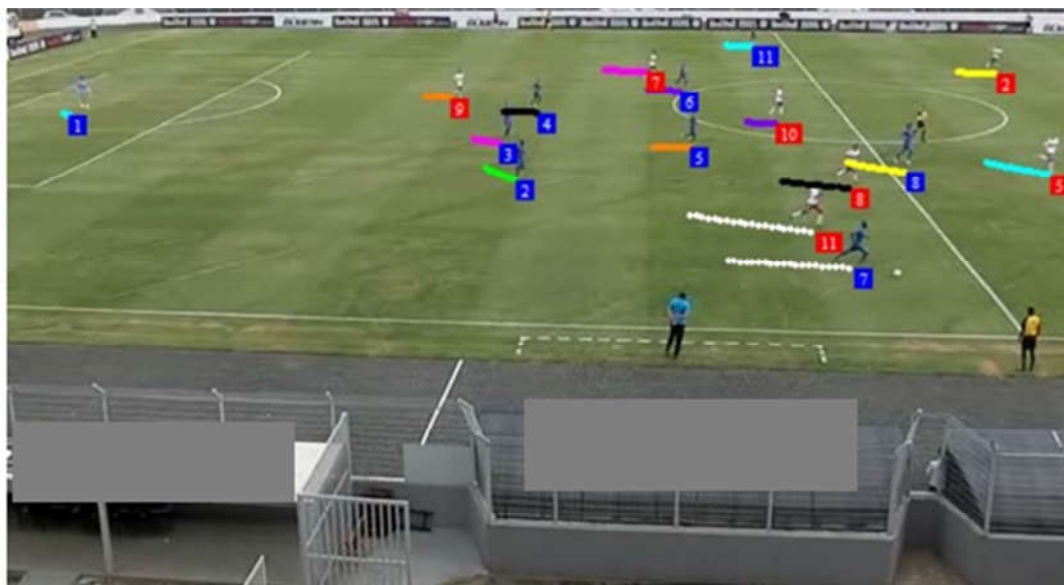


**Figura 3.** Ilustração dos pontos de obtenção das métricas reais no campo (C1, C2,...Cn). (Fonte: o próprio autor).

- Construção dos grafos e rastreamento automático:** As informações obtidas após a segmentação são um conjunto de pontos discretos que podem corresponder à trajetória dos jogadores na sequência de imagens, mas ainda não estão associadas aos respectivos jogadores. Para desenvolver um algoritmo capaz de reconhecer a trajetória de cada jogador em toda sequência de imagens de forma eficiente, a teoria de grafos (Szwarcfiter, 1984) é um método muito utilizado. Um grafo pode ser representado por um conjunto de pontos (vértices) que são conectados por linhas (arestas), ver **Figura 4**. O rastreamento, então, é realizado a partir da construção de grafos utilizando as informações obtidas pelos *blobs*, selecionando o jogador (vértice) no primeiro quadro. Uma vez selecionado o primeiro vértice, o sistema buscará informações sobre o próximo ponto, de acordo com o grafo criado, resultando na identificação do mesmo atleta na sequência de imagens. A **Figura 5** exemplifica o resultado para o rastreamento automático. Posteriormente, a trajetória dos jogadores foi suavizada por meio de um filtro digital Butterworld de 3º ordem, com uma frequência de corte de 0,4 Hz (Misuta et al., 2005; Barros et al., 2007).



**Figura 4.** Sequências de imagens no jogo (A); Obtenção dos blobs correspondente aos jogadores (B); Representação do grafo construído a partir dos blobs (C). (Fonte: o próprio autor).



**Figura 5.** Identificação da posição de cada jogador em campo ao longo de uma sequência de imagens. (Fonte: o próprio autor).

Após a obtenção da posição 2D dos jogadores, foi realizado o registro das ações técnicas realizadas pelos jogadores durante as partidas (**Figura 6**). O software DVideo possui interface para identificar qual ação (ex. chute a gol) um determinado jogador realizou, qual instante (*frame*) do jogo foi realizada esta ação, qual era a posição deste jogador em campo

quando realizou a ação, e se for o caso, identificar se esta ação foi realizada de forma correta ou errada.

Estas informações são importantes neste estudo, para determinar os momentos em que as equipes estavam com e sem posse de bola (Moura et al., 2012) e momentos em que ocorreram finalizações e desarmes.

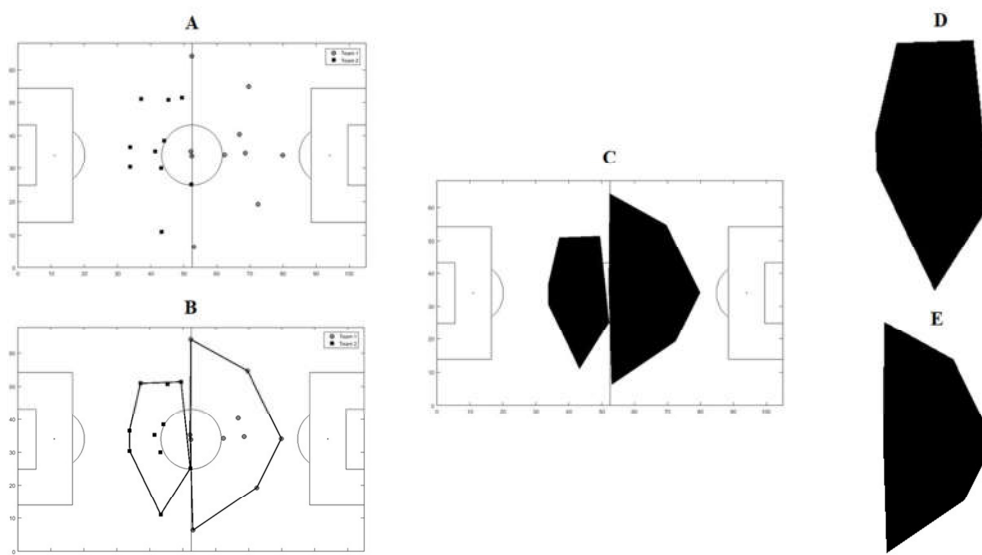


**Figura 6.** Interface para o registro das ações técnicas realizadas pelos jogadores. (Fonte: o próprio autor).

## 4.2 Obtenção das formas de organização das equipes durante as partidas oficiais

Como objeto de estudo a ser analisado, foram geradas imagens de polígonos em ambiente *Matlab* para partidas oficiais dos campeonatos de futebol estaduais e nacionais. Para representar a organização das equipes em campo ao longo do tempo ( $t$ ), foi gerado o envoltório convexo (Preparata e Shamos, 1985; Moura et al., 2012). O envoltório convexo de um conjunto de pontos  $S$  em um plano (neste caso, representado pela posição dos jogadores em campo, sem o goleiro (**Figura 7(A)**), em cada instante de tempo  $t$  é o menor polígono convexo que contém  $S$  (**Figura 7(B)** e **Figura 7(C)**). Se  $S$  é finito, o envoltório convexo é sempre um polígono onde os vértices são um subconjunto de pontos  $S$ . Foi computado o envoltório convexo por meio do método *QuickHull* (Preparata e Shamos, 1985). Finalmente,

para cada equipe, imagens binárias da forma do envoltório convexo (**Figura 7(D)** e **Figura 7(E)**), foram armazenadas a cada instante de tempo.



**Figura 7.** Etapas para a construção das imagens de polígonos convexos que representam a organização das equipes em campo: (a) posição dos jogadores de cada equipe. (b) Aplicação do método *quickhull* para a construção do envoltório convexo. (c) Polígono do envoltório convexo na respectiva posição em campo. Imagens dos formatos analisados para cada equipe, (d) time 1 e (e) time 2. (Fonte: o próprio autor).

### 4.3 Descritores de formas

Para este estudo, foram calculados diferentes descritores de forma para as análises de quais podem melhor representar as imagens dos polígonos analisados, conforme descrito a seguir:

**Dimensão Fractal Multiescala (DFM).** A dimensão fractal (DF) é amplamente utilizada para caracterizar objetos em termos de complexidade, por meio de valores fracionados e sua auto-similaridade (Normant e Tricot, 1991; Plotze et al., 2005). A partir desta definição, a dimensão fractal *Minkowski-Bouligand* tem sido uns dos mais populares em análise de imagens (Torres, Falcão e Costa, 2004). Este método permite a descrição de objetos por meio de dilatações (escalas), por obter um maior número de informações (vetores de características) que pode distinguir objetos em diferentes escalas, sendo este a dimensão fractal multiescala (DFM) (Sarkar e Chaudhuri, 1994; Torres, Falcão e Costa, 2004). A DF de *Minkowski-Bouligand* é calculada de acordo com a (equação 1):

$$DF = 2 - \lim_{r \rightarrow 0} \frac{\log(A(r))}{\log(r)} \quad (1)$$

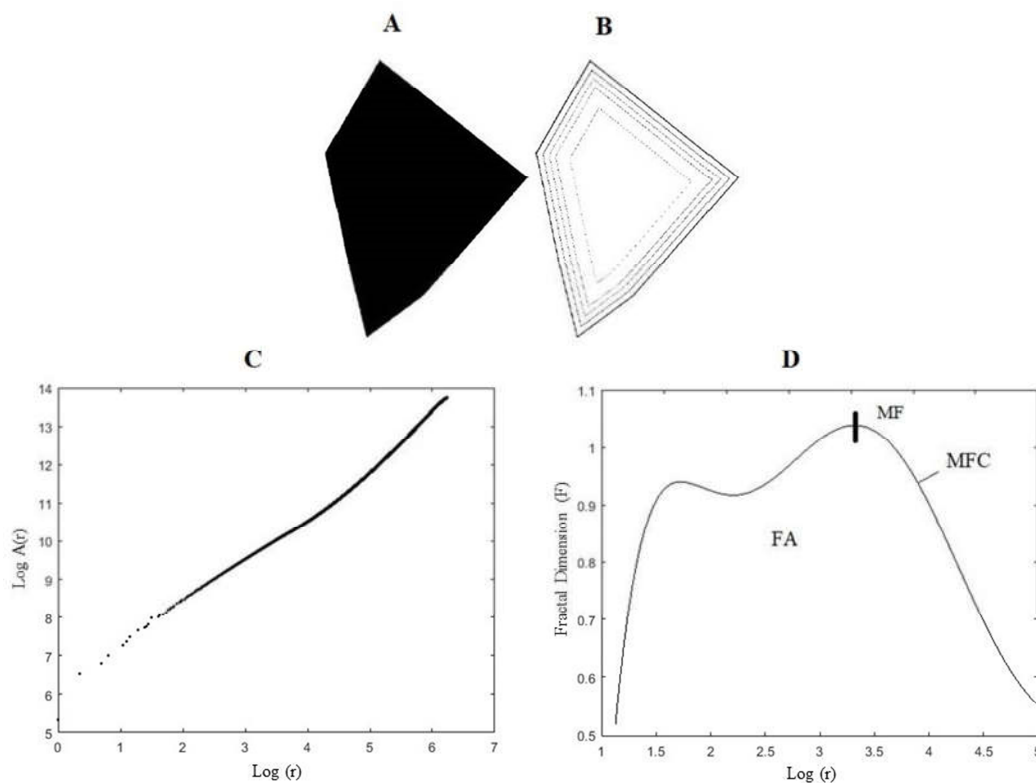
em que  $A(r)$  é a área de uma região dilatada por um raio  $r$ . O algoritmo consiste em três etapas, relatados a seguir, utilizados para calcular a dimensão fractal de *Minkowski-Bouligand* (Torres, Falcão e Costa, 2004).

- Transformação de Distância Euclidiana (TDE) por meio da Image-Foresting Transform:** esta etapa consiste na execução da *Image-Foresting Transform* (representado neste caso pelo mapa de custo) para a realização da TDE das imagens. Dado um conjunto de pontos  $S$  em um plano cartesiano  $(x, y)$ , a dilatação ocorre pela propagação de custo para pixels adjacentes segundo a matriz de distância euclidiana, desde que o novo custo do pixel adjacente seja menor do que o custo anterior (Falcão, Costa e Cunha, 2002). Esta é uma forma de se obter valores progressivos simplificados do formato (**Figura 8(A)** e **Figura 8(B)**). Quanto maiores os valores de distância, maior a simplicidade da forma, conseqüentemente menor os detalhes e similaridades.
- Avaliação das áreas de contornos dilatados:** cada instância multiescala  $S(r)$  da forma original  $S$  é obtida pelo limiar do mapa de custo em uma determinada distância euclidiana  $r$  ao quadrado; dessa forma é possível obtermos a área de uma instância multiescala,  $A(r)$ , por meio do histograma cumulativo do mapa de custos. Da equação 1, também é necessário computar o  $\log x \log$  destes histogramas cumulativos.
- Estimativa da dimensão fractal multiescala:** definida como uma abordagem comum para avaliar uma dimensão fractal com um único valor ajustável linearmente a curva de  $\log A(r) \times \log r$ , e considera a dimensão fractal  $F$  como 2 menos o coeficiente angular. Para este estudo, utilizou-se um método proposto previamente (Torres, Falcão e Costa, 2004), cuja uma curva com um polinômio  $f_n(r)$  com  $n$  graus é ajustada, por meio de regressão, à função logarítmica da área a partir da qual as derivadas podem ser obtidas imediatamente. Assim, a dimensão fractal dos contornos dilatados é obtida em função dos raios de dilatação, (equação 2):

$$F(r) = 2 - f'_n(r) \quad (2)$$

A **Figura 8(C)** apresenta o resultado do histograma cumulativo de  $\log x \log$  do mapa de custo filtrado com um polinômio de grau  $n$  ( $n = 10$ ) para um vetor de características com amostra de 100 pontos em um exemplo mostrado na **Figura 8(B)**. Finalmente, a **Figura 8(D)**, apresenta o resultado da DFM para a **Figura 8(A)**.

Depois que todas as etapas descritas acima foram realizadas, obtém-se a curva fractal multiescala (CFM) para cada imagem da base de dados. A CFM consiste nos valores de dimensão fractal escalar variante de um determinado polígono. Para caracterizar a forma das figuras, também foi identificado o máximo valor da CFM (MF), e também, foi calculada a integral da CFM, obtendo-se os valores de área (AF) abaixo da CFM.



**Figura 8.** (a) Imagem de um polígono; (b) Função logarítmica da área; (c) Dimensão Fractal Multiescala da imagem analisada (Máximo Valor Fractal (MF); Curva Fractal Multiescala (CFM), área fractal (AF)). (Fonte: o próprio autor).

**Retangularidade.** A retangularidade pode ser definida como uma medida do quão bem uma forma pode aproximar-se ao menor retângulo que delimita o polígono (MER). A retangularidade é calculada de acordo com a equação 3:

$$\text{Retangularidade} = \frac{A(\text{shape})}{A(\text{MER})} \quad (3)$$

em que  $A(\text{shape})$  representa a área do polígono (em *pixels*) e  $A(\text{MER})$  representa a área do menor retângulo que delimita o polígono (em *pixels quadrados*). Esta medida fornece resultados de 0 a 1; quanto mais próximo de 1, mais retangular é a forma.

**Circularidade.** Neste estudo, a circularidade foi definida como a razão entre o perímetro ao quadrado e a área (equação 4):

$$\text{Circularidade} = \frac{P^2}{A} \quad (4)$$





em que  $P$  e  $A$  denota o perímetro e a área do polígono, respectivamente. A circularidade fornece uma medida de quanto mais próxima a forma é circular, o valor tende a zero. Ambas as medidas (retangularidade e circularidade) são adimensionais (Costa e Cesar Jr, 2009).

**Relação Área/Perímetro.** Descritor de forma que demonstra a relação Área/Perímetro (AP) também foi calculado, definido de acordo com a (equação 5):

$$AP = \frac{A}{P} \quad (5)$$

em que  $A$  e  $P$  denota a área e o perímetro do polígono, respectivamente. Exemplos de valores destes descritores para formas genéricas são apresentados no **Quadro 1**.

**Quadro 1.** Cálculo de descritores de formas Retangularidade, Circularidade e Área/Perímetro para exemplos de formas genéricas.

Descritores de formas				
Retangularidade	1,00	0,78	0,53	0,75
Circularidade	15,51	12,64	24,61	13,57
Área/Perímetro	36,81	36,33	23,85	38,37

*Alldescriptors*. Finalmente, um vetor de características (VC) foi criado por meio da concatenação dos valores de todos os descritores, denominado *Alldescriptors* (AllD). Nesse VC, os dados de CFM não foi incluso. O mesmo ficou organizado da seguinte forma:

Alldescriptors = [AP AF Circularidade MF Retangularidade];

## 5 Estudo um: Dimensão fractal multiescala aplicada às análises táticas no futebol: uma nova abordagem para avaliar as formas de organização de equipas em campo<sup>2</sup>.

### Resumo

O objetivo deste estudo foi avaliar diferentes descritores de formas aplicados às imagens de polígonos que representam a organização de equipas de futebol em campo. A eficácia de diferentes descritores de forma (área/perímetro, área fractal, circularidade, máximo fractal, retangularidade, curva fractal multiescala – CFM), e a concatenação de todos os descritores de formas (exceto CFM), denominado *All descriptors* (AllID) foram avaliados e aplicados a polígonos correspondentes a formas, representadas pelo envoltório convexo obtidos pelas coordenadas 2D dos jogadores. Um sistema de recuperação de imagens baseado em conteúdo (SRIBC) foi aplicado para 25 usuários (idade média de  $31,9 \pm 8,4$  anos) para avaliar a relevância de imagens. Medidas de eficácia foram utilizadas para avaliar os descritores de formas ( $P@n$  and  $R@n$ ). Os descritores CFM ( $P@5 = 0,46 \pm 0,37$  e  $P@10 = 0,40 \pm 0,31$ ,  $p < 0.001$ ;  $R@5 = 0,14 \pm 0,13$  e  $R@10 = 0,24 \pm 0,19$ ,  $p < 0.001$ ) e AllID ( $P@5 = 0,43 \pm 0,36$  e  $P@10 = 0,39 \pm 0,32$ ,  $p < 0.001$ ;  $R@5 = 0,13 \pm 0,11$  e  $R@10 = 0,24 \pm 0,20$ ,  $p < 0.001$ ) apresentaram altos valores de eficácia. Como uma demonstração prática, o melhor descritor de formas avaliado (CFM) foi aplicado para análise tática em uma partida oficial. A técnica de cluster *k-means* foi aplicada e diferentes formas de organização podem ser identificadas ao longo da partida. A CFM é o mais eficaz descritor de formas em relação aos outros avaliados, e torna-se possível aplicar este descritor em análises de partidas de futebol profissional.

**Palavras-chave:** esportes coletivos; análise de partidas; tática; dimensão fractal; recuperação de imagens baseado em conteúdo.

---

<sup>2</sup> Bueno, M. J. O., et al. Multiscale fractal dimension applied to tactical analysis in football: A novel approach to evaluate the shapes of team organization on the pitch. *Plos One*, v.16, n.9, p.1-14. 2021.

## 5.1 Introdução

Com a possibilidade de obtenção de dados sobre a posição dos atletas em função do tempo, analisar a formação de jogadores da equipe e adversários, assim como padrões de jogo, tem sido objetivo de pesquisas recentes sobre futebol (Frencken et al., 2011; Moura et al., 2012; Sarmiento et al., 2017; Memmert et al., 2019; Low et al., 2020). Vários métodos computacionais têm sido desenvolvidos e aplicados para compreender o comportamento tático em esportes coletivos relacionados a parâmetros espaciais (Rico-González et al., 2020(a); Rico-González et al., 2020(b); Rico-González et al., 2021). Estas variáveis são apresentadas na literatura como componentes relevantes para as análises na organização de equipes de futebol em campo, como centroide, espalhamento, *stretch index*, largura, comprimento e área de ocupação que as equipes ocupam em campo (Bartlett et al., 2012; Moura et al., 2012; Duarte R. et al., 2013; Clemente M. F. et al., 2017). Na literatura, a área de ocupação é normalmente definida como a área do menor polígono convexo formado pela posição dos jogadores, representando o total de espaço ocupado por uma equipe em campo (Frencken e Lemmink, 2009; Moura et al., 2012).

Apesar de os valores de magnitude da área de ocupação serem considerados bons parâmetros para a caracterização da organização de equipes durante uma partida e amplamente utilizados (Frencken et al., 2011; Bartlett et al., 2012; Moura et al., 2012; Vilar et al., 2013), pode-se argumentar que diferentes distribuições de jogadores em campo podem fornecer áreas de ocupação iguais, ou ainda, para uma mesma forma de organização, é possível que equipes apresentem diferentes áreas de ocupação. Assim, somente pela magnitude da área de ocupação ( $m^2$ ), não é possível estabelecer em detalhes como as equipes estão organizadas. Nesse sentido, uma análise sobre a forma estrutural do polígono, representada pela área de ocupação de equipes de futebol em campo, pode fornecer informações sobre a complexidade da organização de equipes durante partidas oficiais.

A forma estrutural de um objeto é considerada uma importante característica útil para descrever e distinguir determinados objetos (Locaric, 1998). Alguns descritores de formas são apresentados na literatura para a caracterização de um determinado objeto, tal como a circularidade, retangularidade, e comprimento do maior e menor eixo da imagem (eixos principais) (Costa e Cesar Jr, 2009). Estes descritores são classificados basicamente em duas categorias: análise baseada em contorno ou baseada em região (Zhang e Lu, 2004). Como

bem rotulados, a primeira usa apenas informações do contorno do objeto, enquanto a segunda emprega todos os *pixels* de um objeto para obter um vetor de características.

Uma abordagem eficaz para estimar a complexidade de uma determinada forma consiste no uso de medidas de dimensão fractal. A geometria fractal é um campo da matemática frequentemente empregada para analisar a complexidade de uma forma (Mandelbrot, 1982). Esta teoria tem sido utilizada em processamento de imagens, especificamente para segmentação de texturas (Chaudhuri e Sarkar, 1995) e compressão de imagens (Fisher, 1995). Uma destas definições fundamentais da dimensão fractal é a medida de complexidade que generaliza o conceito de dimensão topológica. A Dimensão Fractal Multiescala (DFM), é uma extensão da dimensão fractal *Minkowski-Bouligand*, que utiliza a dilatação morfológica Euclidiana para representação multiescala, como característica a auto-similaridade (Torres, Falcão e Costa, 2004).

Uma vez que ferramentas que descrevem as propriedades de formas de polígonos associadas à organização de equipes de futebol não são apresentadas na literatura, fazem-se necessárias a implementação, avaliação e validação de métodos de descritores de formas, que possam melhor descrever esses polígonos. Portanto, o objetivo deste estudo foi avaliar diferentes descritores de formas aplicadas às imagens de polígonos que representam a organização de equipes de futebol em campo, e estabelecer qual(is) o(s) melhor(es) descritor(es) para descrever as formas de organização de equipes ao longo de uma partida. A hipótese inicial é que a DFM pode apresentar melhores resultados para descrever a forma de organização de equipes de futebol, de acordo com o fato de que o DFM não é invariante à escala (Torres, Falcão e Costa, 2004), o que permite a identificação de formas com complexidade semelhante, considerando a respectiva escala (diferentes áreas, por exemplo). Adicionalmente, demonstramos uma possível aplicação do descritor mais bem avaliado para descrever a organização de equipes de futebol ao longo de uma partida oficial.

## **5.2 Materiais e métodos**

### **5.2.1 Coleta de dados**

Para a obtenção da área de ocupação e obtenção dos polígonos correspondentes à forma de organização das equipes em campo ao longo do tempo de uma partida, a posição 2D

dos jogadores em campo foi obtida por meio do *software DVideo* (Figuroa, Leite e Barros, 2006a; 2006b). O *software* possui uma média de erro de 0,3 m para a determinação da posição do jogador e uma média de erro de 1,4% para a distância percorrida pelos jogadores (Figuroa, Leite e Barros, 2006b; Misuta, 2009). O protocolo deste estudo foi aprovado e conduzido de acordo com as diretrizes éticas estabelecidas pela instituição local – Comitê de ética da Universidade Estadual de Londrina (parecer nº 3.047.461) e em conformidade com as recomendações da Declaração de *Helsinki*.

### **5.2.2 Elaboração dos polígonos que representam a organização tática de equipes de futebol.**

Como objeto de estudo a ser analisado, foram geradas imagens de polígonos em ambiente *Matlab* para uma partida oficial do campeonato brasileiro de futebol da primeira divisão, de ambas as equipes envolvidas. Para representar a organização das equipes em campo ao longo do tempo ( $t$ ), foi gerado o envoltório convexo (Preparata e Shamos, 1985; Moura et al., 2012). O envoltório convexo de um conjunto de pontos  $S$  em um plano (neste caso, representado pela posição dos jogadores em campo, sem o goleiro, em cada instante de tempo  $t$  é o menor polígono convexo que contém  $S$ . Se  $S$  é finito, o envoltório convexo é sempre um polígono onde os vértices são um subconjunto de pontos  $S$ . Foi computado o envoltório convexo por meio do método *QuickHull* (Preparata e Shamos, 1985). Finalmente, para cada equipe, imagens binárias da forma do envoltório convexo, foram armazenadas a cada instante de tempo.

### **5.2.3 Descritores de formas**

Para este estudo, foram calculados diferentes descritores de forma para as análises de quais podem representar melhor as imagens dos polígonos analisados, conforme descrito brevemente abaixo (a descrição completa de todos os descritores pode ser vista na seção 4.3):

**Dimensão Fractal Multiescala (DFM).** A dimensão fractal (DF) é amplamente utilizada para caracterizar objetos em termos de complexidade, por meio de valores fracionados e sua auto-similaridade (Normant e Tricot, 1991; Plotze et al., 2005). A partir desta definição, a dimensão fractal *Minkowski-Bouligand* tem sido um dos mais populares em análise de imagens (Torres, Falcão e Costa, 2004). Este método permite a descrição de objetos

por meio de dilatações (escalas), por obter um maior número de informações (vetores de características) que pode distinguir objetos em diferentes escalas, sendo este a dimensão fractal multiescala (DFM). O algoritmo consiste em três etapas, relatadas a seguir, utilizadas para calcular a dimensão fractal de *Minkowski-Bouligand* (Torres, Falcão e Costa, 2004).

- **Transformação de Distância Euclidiana (TDE) por meio da Image-Foresting Transform:** esta etapa consiste na execução da *Image-Foresting Transform* (representada neste caso pelo mapa de custo) para a realização da TDE das imagens. Dado um conjunto de pontos  $S$  em um plano cartesiano  $(x, y)$ , a dilatação ocorre pela propagação de custo para *pixels* adjacentes segundo a matriz de distância euclidiana, desde que o novo custo do *pixel* adjacente seja menor do que o custo anterior (Falcão, Costa e Cunha, 2002).
- **Avaliação das áreas de contornos dilatados:** cada instância multiescala  $S(r)$  da forma original  $S$  é obtida pelo limiar do mapa de custo em uma determinada distância euclidiana  $r$  ao quadrado; dessa forma é possível obter a área de uma instância multiescala,  $A(r)$ , por meio do histograma cumulativo do mapa de custos.
- **Estimativa da dimensão fractal multiescala:** definida como uma abordagem comum para avaliar uma dimensão fractal com um único valor ajustável linearmente a curva de  $\log A(r) \times \log r$ , e considera a dimensão fractal  $F$  como 2 menos o coeficiente angular (Torres, Falcão e Costa, 2004). O histograma cumulativo de  $\log x \log$  do mapa de custo foi filtrado com um polinômio de grau 10 para um vetor de características com amostra de 100 pontos em um exemplo mostrado na **Figura 8(B)**. Finalmente, a **Figura 8(D)**, apresenta o resultado da DFM para a **Figura 8(A)**.

Depois que todas as etapas descritas acima foram realizadas, obtém-se a curva fractal multiescala (CFM) para cada imagem da base de dados. A CFM consiste nos valores de dimensão fractal escalar variante de um determinado polígono. Para caracterizar a forma das figuras, também foi identificado o máximo valor da CFM (MF), e também, foi calculada a integral da CFM, obtendo-se os valores de área (AF) abaixo da CFM.

**Retangularidade:** A retangularidade pode ser definida como uma medida do quão bem uma forma pode aproximar-se ao menor retângulo que delimita o polígono (MER).

**Circularidade:** Nesse estudo, a circularidade foi definida como a razão entre o perímetro ao quadrado e a área. Ambas as medidas (retangularidade e circularidade) são adimensionais (Costa e Cesar Jr, 2009).

**Relação Área/Perímetro:** Descritor de forma que demonstra a relação Área/Perímetro (AP) também foi calculado.

**Alldescriptors:** Finalmente, um vetor de características (VC) foi criado por meio da concatenação dos valores de todos os descritores, denominado *Alldescriptors* (AllD). Nesse VC, os dados de CFM não foram inclusos. O mesmo ficou organizado da seguinte forma:

Alldescriptors = [AP AF Circularidade MF Retangularidade];

De posse dos valores de descritores de formas calculados para cada polígono em função do tempo, foi implementado um sistema para avaliar os descritores e identificar polígonos mais relevantes em relação às imagens de consulta (*query*) selecionadas aleatoriamente na mesma base de dados. Para tanto, foi utilizado um sistema computacional baseado em Recuperação de Imagens por Conteúdo (Torres e Falcão, 2006) que será descrito na seção seguinte.

#### **5.2.4 Sistema de Recuperação de Imagens Baseado em Conteúdo (SRIBC): etapas para a implementação**

Para a implementação de um Sistema de Recuperação de Imagens por Conteúdo (SRIC), o descritor de imagens é muito importante, este sendo importante para avaliar a similaridade de imagens de consultas (Torres e Falcão, 2006). Neste estudo foram selecionadas 25 imagens de consulta (*queries*) aleatoriamente na mesma coleção de imagens do jogo em questão. Estes *queries* representam diferentes polígonos que podem ser extraídos de uma mesma base de dados, o que torna possível explorar uma grande variedade de formas

que uma equipe pode apresentar ao longo de uma partida. O processo para a elaboração de um SRIBC ocorre na seguinte ordem (Penatti, Valle e Torres, 2012):

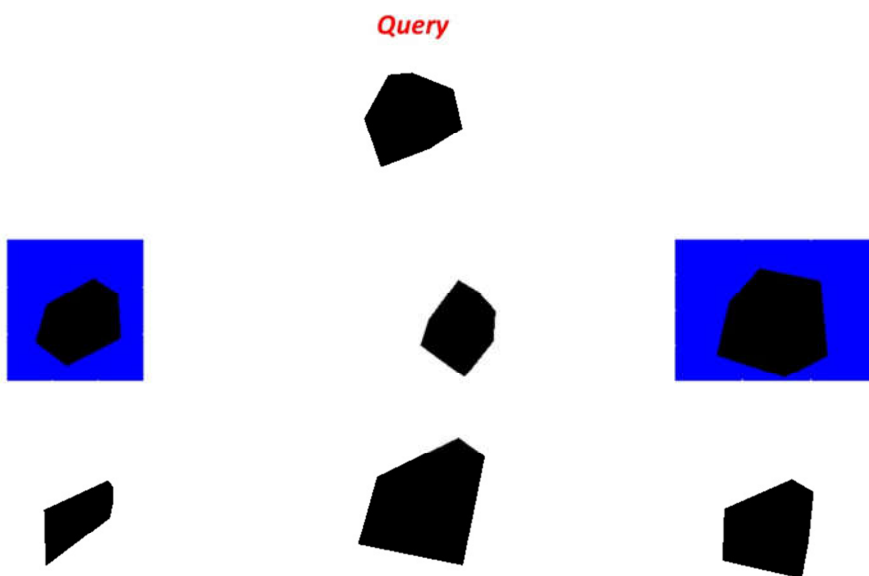
- Os descritores de formas das *queries* foram calculados, assim como para todas as imagens da coleção. Este passo é importante para busca de imagem baseada em conteúdo. Posteriormente, uma análise de similaridade foi realizada entre as *queries* e a coleção de imagens.
- As análises de similaridade entre as *queries* e as imagens da coleção foram calculadas por meio da distância euclidiana (em computação denominada *L2*) entre os valores dos VC de cada uma das *queries* para todas as imagens da coleção.
- Um ranqueamento das imagens da coleção foi realizado baseado na distância Euclidiana, do mais similar (menor distância) para o menos similar (maior distância) às *queries*. Esta etapa foi realizada para que o SRIBC possa obter as imagens mais relevantes de cada descritor para cada imagem consulta. Um descritor é considerado melhor do que outro se, ao ser utilizado, leva a mais resultados relevantes nas primeiras posições do ranking.

Após a realização de todas as etapas para a construção de um SRIBC, as imagens relevantes (imagens similares), obtidas pelo SRIBC, foram avaliadas por usuários reais, de acordo com as orientações prévias da literatura (Gonçalves et al., 2004; Penatti, Valle e Torres, 2012). Para cada descritor analisado, foram recuperadas as 12 primeiras imagens mais relevantes de cada imagem de consulta. Para cada lista de imagens selecionadas em cada modelo, essas imagens foram unidas em uma única lista e em seguida foram removidas as duplicadas. Ao final, as listas de imagens foram embaralhadas antes de serem exibidas aos usuários por meio de uma interface criada especificamente para este estudo.

### 5.2.5 Avaliação da eficácia de descritores de formas.

**Participantes:** Para esse estudo, 25 usuários (estudantes e profissionais em ciências dos esportes), de ambos os sexos, idade média de  $31,9 \pm 8,4$  anos, foram convidados a avaliar imagens relevantes obtidas pelo SRIBC. Os usuários visitaram o laboratório em uma única sessão e assinaram o termo de consentimento livre e esclarecido (apêndice II) aprovado pelo Comitê de Ética em Pesquisa de uma universidade local (processo nº 3.047.461) e conduzido de acordo com as recomendações éticas da Declaração de *Helsinki*.

**Interface para avaliação de relevância das imagens.** Foi criada uma interface no *software MATLAB* na qual era apresentada uma dada imagem de consulta (*query*) em destaque aos participantes e, abaixo, as imagens a serem analisadas (**Figura 9**). O usuário, por meio da interface, teve que analisar seis imagens e identificar quais entre os seis polígonos eram mais relevantes em relação à *query*. Após identificar a(s) imagem(ns) que considerou relevante(s), o mesmo clicava na imagem e esta era rotulada (**Figura 9**, Azul). Passava-se então para outras seis imagens que poderiam ser relevantes para a mesma *query*. Foi instruído aos participantes para que avaliassem para três imagens consultas, 84 imagens possivelmente relevantes, sendo, portanto, um total de 252 imagens analisadas. Assim que cada participante encerrava a avaliação, uma matriz de dados era gerada e salva com as informações de quais imagens o participante considerou relevantes para cada consulta.



**Figura 9.** Exemplo da interface para a avaliação da relevância dos polígonos em relação ao *Query* (imagem consulta). Abaixo os 6 polígonos para que o usuário possa avaliar sua relevância. Em azul um exemplo de polígonos possivelmente relevantes identificados pelo usuário. (Fonte: o próprio autor).

**Avaliação do SRIBC.** Medidas de *precision* e *recall* foram calculadas para avaliar a eficácia das listas ranqueadas retornadas por diferentes descritores. A *precision* mede a fração de imagens relevantes retornadas pelo usuário em uma determinada consulta (um polígono de entrada) em relação ao total de imagens. A *recall* mede a fração de imagens relevantes

retornadas para uma determinada consulta em relação ao total de imagens relevantes existentes na base (Penatti, Valle e Torres, 2012).

A *precision* na *n-ésima* posição ( $P@n$ ) e a *recall* para *n-ésima* posição ( $R@n$ ) foram calculadas para um determinado número de imagens, que fornece uma avaliação da impressão do usuário sobre os resultados ranqueados.  $P@5$  e  $P@10$  fornecem métricas confiáveis para avaliar se máquinas de busca obtêm documentos relevantes no topo do ranking. Assim, quanto maior a concentração de imagens relevantes no topo do ranking, mais eficaz é considerado o descritor avaliado (Baeza-Yates e Ribeiro-Neto, 1999).

A avaliação de eficácia do sistema foi medida por meio de curva de *precision x recall* e por valores de  $P@5$ ,  $P@10$ ,  $R@5$  e  $R@10$ . Abaixo são apresentados os cálculos:

$$Precision = \frac{r}{n}$$

$$Recall = \frac{r}{rt}$$

em que  $r$  é o número de imagens relevantes para o descritor,  $n$  o número total de imagens na coleção e  $rt$  o número total de imagens consideradas relevantes para cada consulta.

### 5.2.6 Análises estatísticas.

Uma medida de concordância interpessoal foi testada pelo cálculo de coeficiente de Kappa (Cohen, 1960), para verificar a concordância entre diferentes avaliadores em identificar a relevância das imagens para as consultas, sendo que para cada query, 3 indivíduos fizeram a avaliação, denominados A, B e C. As análises de concordância foram realizadas par a par (A x B; A x C; e B x C). Foram considerados os seguintes valores de referências:  $k \leq 0.2$  considerado leve;  $k > 0.2 - 0.4$  como razoáveis;  $k > 0.4 - 0.6$  como moderado;  $k > 0.6 - 0.8$  considerável;  $k > 0.8$  como quase perfeito (Landis e Koch, 1977).

Um teste de *Levene* foi realizado para a análise de homocedasticidade dos dados. Uma análise de esfericidade foi aplicada por meio do teste de *Mauchly*. Uma *anova one-way* de medidas repetidas foi aplicada para verificar se houve diferenças entre os descritores para os valores de  $P@5$ ,  $P@10$  e  $R@5$  e  $R@10$ . Quando diferenças foram encontradas, um teste de *post-hoc* de *Bonferroni* foi aplicado para fornecer informações específicas em quais dados

foram diferentes uns dos outros. Para todas as análises, foi adotado um valor de significância de  $p < 0,05$ . Todos os resultados foram apresentados em média  $\pm$  desvio-padrão.

### 5.2.7 Resultados

A **Tabela 1** apresenta os resultados de concordância entre os avaliadores para a identificação dos formatos relevantes durante o processo de avaliação de eficácia. É possível observar uma concordância moderada entre os avaliadores A x B e A x C, quanto à identificação de formatos relevantes a consultas em relação ao descritor CFM. Entre B X C, apresentou maiores valores para ‘Alldescriptors’. A **Tabela 2** apresenta resultados de percentual quando os avaliadores consideraram as imagens não relevantes (NR) e relevantes (R).

**Tabela 1.** Resultados das análises de concordância (Coeficiente de Kappa) para identificação de relevância dos formatos para cada descritor entre os 3 diferentes avaliadores (A, B, e C).

	A x B	A x C	B x C
Área/Perímetro	0,22	0,19	0,31
Área Fractal	0,35	0,24	0,21
Circularidade	0,21	0,20	0,23
Máximo Fractal	0,14	0,07	0,16
Retangularidade	0,15	0,24	0,26
CFM	0,49	0,45	0,30
Alldescriptors	0,35	0,39	0,54

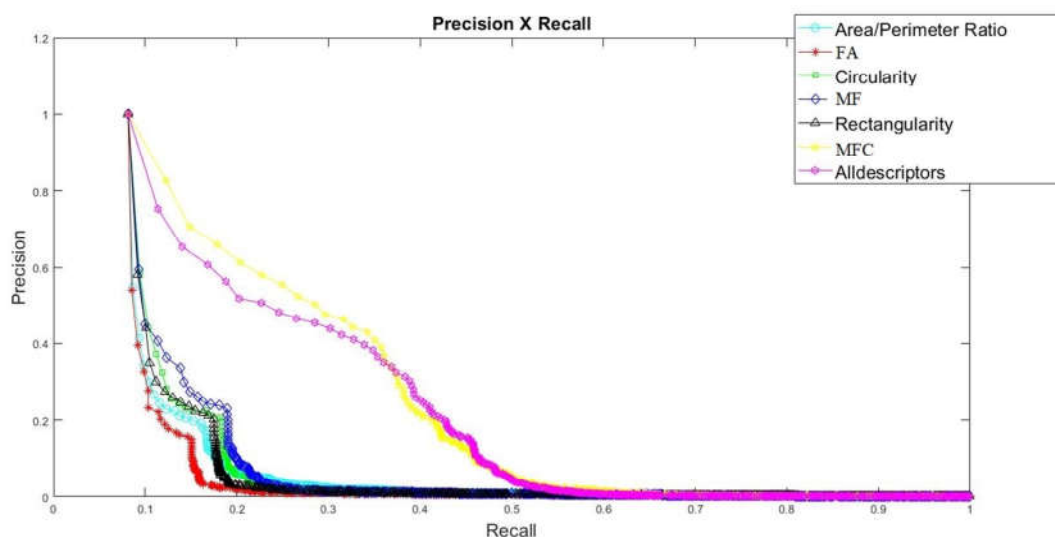
Abreviações: CFM = Curva Fractal Multiescala;

**Tabela 2.** Resultados do percentual de concordância entre imagens consideradas não relevantes (NR - NR) e imagens consideradas relevantes (R - R) entre 3 diferentes avaliadores.

	A x B		A x C		B x C	
	NR - NR	R - R	NR - NR	R - R	NR - NR	R - R
Área/Perímetro	76,66	4,66	80,00	3,33	79,66	5,00
Área Fractal	82,66	4,66	85,33	2,66	85,66	7,33
Circularidade	78,33	4,00	74,66	5,00	77,00	4,66
Máximo Fractal	72,66	4,33	75,66	2,66	70,66	5,33
Retangularidade	77,66	3,33	79,00	4,33	76,66	5,33
CFM	52,00	24,00	57,33	17,00	45,66	21,00
Alldescriptors	50,00	20,33	54,00	19,33	51,00	27,33

Abreviações: CFM = Curva Fractal Multiescala;

A **Figura 10** ilustra o desempenho de busca, em termos de *precision x recall*, para toda a base de dados ranqueada dos 7 descritores avaliados. Pode-se observar que descritores que obtêm um maior número de características de informações (vetores de características), obtiveram melhores resultados de eficácia, sendo estes curva fractal multiescalar (CFM) e em seguida Alldescriptors (AllD), resultados esses que podem ser confirmados na **Tabela 3**. Diferenças estatísticas foram observados entre os descritores para P@5 ( $F = 33,62$ ;  $p < 0,001$ ) e para P@10 ( $F = 32,15$ ;  $p < 0,001$ ). O teste *post-hoc* mostrou para P@5 ( $F = 14,62$ ;  $p < 0,001$ ) que os resultados dos descritores CFM e AllD apresentaram valores de precisão estatisticamente superiores em relação aos demais descritores. Para P@10 ( $F = 11,76$ ;  $p < 0,001$ ), CFM e AllD também apresentaram valores de precisão estatisticamente maiores em comparação com todos os outros descritores. O descritor FA apresentou diferenças estatísticas para MF em P@5 e P@10. Para R@5, CFM e AllD apresentaram valores estatisticamente superiores em relação a todos os outros descritores ( $F = 21,45$ ;  $p < 0,001$ ). O descritor FA apresentou valores menores entre os MF. Para R@10, CFM e AllD, os resultados também foram maiores em relação a todos os outros descritores ( $F = 22,26$ ;  $p < 0,001$ ).



**Figura 10.** Representação dos resultados de desempenho de busca em relação a curva Precision x Recall para todos os descritores avaliados. (Fonte: o próprio autor).

**Tabela 3.** Valores de média  $\pm$  desvio padrão para precision (P@5 e P@10) e recall (R@5 e R@10) para avaliar a eficácia de diferentes descritores.

	Precision (P@5)	Recall (R@5)	Precision (P@10)	Recall (R@10)
Área/Perímetro	0,12 $\pm$ 0,22	0,03 $\pm$ 0,05	0,13 $\pm$ 0,17	0,07 $\pm$ 0,09
Área Fractal	0,09 $\pm$ 0,13 <sup>†</sup>	0,03 $\pm$ 0,04 <sup>†</sup>	0,09 $\pm$ 0,13 <sup>†</sup>	0,06 $\pm$ 0,10
Circularidade	0,14 $\pm$ 0,18	0,04 $\pm$ 0,08	0,14 $\pm$ 0,16	0,09 $\pm$ 0,10
Máximo Fractal	0,20 $\pm$ 0,20	0,06 $\pm$ 0,08	0,17 $\pm$ 0,15	0,10 $\pm$ 0,09
Ratangularidade	0,14 $\pm$ 0,17	0,04 $\pm$ 0,06	0,14 $\pm$ 0,14	0,09 $\pm$ 0,09
CFM	0,46 $\pm$ 0,37 <sup>#*†a</sup>	0,14 $\pm$ 0,13 <sup>#*†a</sup>	0,40 $\pm$ 0,31 <sup>#*†a</sup>	0,24 $\pm$ 0,19 <sup>#*†a</sup>
AllDescriptors	0,43 $\pm$ 0,36 <sup>#*†a</sup>	0,13 $\pm$ 0,11 <sup>#*†a</sup>	0,39 $\pm$ 0,32 <sup>#*†a</sup>	0,24 $\pm$ 0,20 <sup>#*†a</sup>

<sup>#</sup>: Significativamente diferente de área/perímetro;

<sup>\*</sup>: Significativamente diferente da área fractal;

<sup>†</sup>: Significativamente diferente da circularidade;

<sup>‡</sup>: Significativamente diferente do máximo fractal;

<sup>a</sup>: Significativamente diferente da retangularidade;

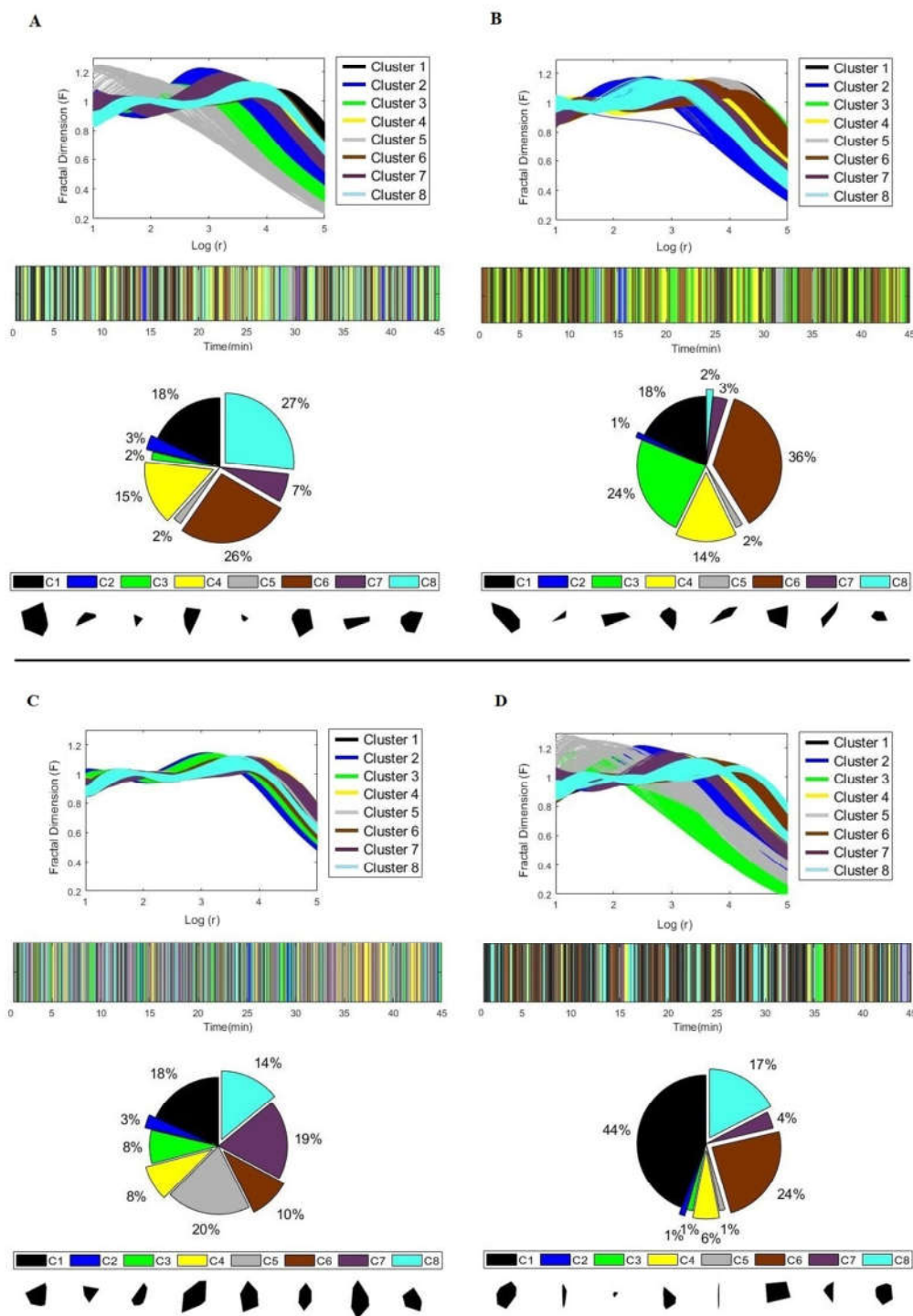
Abreviações: CFM = Curva Fractal Multiescal;

### 5.2.8 Exemplo de aplicação utilizando descritores de formas para análises de partida no futebol.

**Coleta de dados.** Para fins de exemplificação, foi realizado uma aplicação do método em uma partida de futebol profissional brasileiro, com aplicação do melhor descritor avaliado. Foi obtida a posição 2D dos jogadores por meio de um sistema baseado em vídeo (seção 4.1) e obtidas as imagens das formas que representam a organização das equipes (seção 4.2) ao longo da partida para o primeiro e segundo período de ambas as equipes (Equipe 1; Equipe 2). Em seguida, foi realizado o cálculo do descritor de forma CFM para a descrição das imagens (seção 4.3).

**Análise dos dados.** De posse dos valores da CFM para cada imagem, um método de cluster *k-means* (Kaufman e Rousseeuw, 1990), foi aplicado para agrupar os dados de CFM e poder identificar quais as formas que as equipes adotaram durante a partida. Este agrupamento foi realizado por meio dos dados de Distância Euclidiana (DE) dos vetores de características das CFM. O método *Elbow* (Han, Kamber e Pei, 2012) foi utilizado para identificar qual era número de clusters *k* ideal a ser analisado, com o resultado de  $k = 8$ . Depois de identificado quais as formas foram realizadas durante a partida, foi quantificado o número de formas em cada cluster e, então, calculado o percentual de ocorrência na partida.

**Resultados.** A Fig. 12 mostra os resultados de agrupamento para a CFM da Equipe 1 (**Figura 11A**, 1° período; **Figura 11B**, 2° período) e Equipe 2 (**Figura 11C**, 1° período; **Figura 11D**, 2° período) nos 8 clusters. Com o agrupamento da CFM, foi possível identificar quais formas as equipes atuam durante a partida. A **Figura 11** também apresenta os resultados, com exemplos de quais formas podem ser observadas durante a partida, quando essas formas ocorrem durante a partida (gráfico de barras) e a porcentagem de formas (gráfico de pizza) para cada cluster (C1 ,..., C8 ) da Equipe 1 (**Figura 11A**, 1° período; **Figura 11B**, 2° período) e Equipe 2 (**Figura 11C**, 1° período; **Figura 11D**, 2° período).



**Figura 11.** Resultado do agrupamento da CFM para as equipes 1 e 2: (A) e (B), 1° e 2° período da equipe 1, respectivamente; (C) e (D), 1° e 2° período da equipe 2, respectivamente. Resultados também das formas encontradas para cada cluster analisado, quando ocorrem ao longo da partida (gráfico de barras) e percentis de formas em cada cluster (gráfico de pizza). (Fonte: o próprio autor).

### 5.3 Discussão

O objetivo do presente estudo foi realizar uma análise comparativa de diferentes descritores de formas e assim saber qual(is) descritor(es) é(são) melhor(es) para caracterizar os formatos de organização de equipes de futebol brasileiras. Também foi apresentado neste estudo um exemplo de aplicação prática do melhor descritor de formas avaliado, para identificar quais as formas nas quais as equipes de futebol se organizam em campo durante a partida.

Criar um sistema de busca de imagens baseado por conteúdo não é uma tarefa simples, uma vez que requer a criação de algoritmos de codificação de informações de uma imagem em vetores de características (VC) e medidas de similaridade para comparação de uma determinada imagem consulta para uma base de dados (Torres e Falcão, 2006), sendo comumente avaliados por meio de usuários reais (Mokhtarian, Abbasi e Kittler, 1996; Abbasi, Mokhtarian e Kittler, 2000).

Os resultados podem ser observados na **Figura 10**, sendo apresentado o desempenho dos descritores em relação às medidas de *precision x recall*. Pode-se observar que o descritor CFM apresentou um melhor desempenho para a busca de imagens em relação a outros descritores avaliados, como também o descritor Alldescriptors apresentou um bom desempenho. Porém, para a obtenção dos valores do VC Alldescriptors das imagens na base de dados, foi necessário um custo computacional maior para calcular todos os descritores e organizar em um VC, o que torna o descritor Alldescriptors menos eficiente em relação ao tempo de execução. Mesmo assim, é possível destacar que descritores com o maior número de informações (vetor de característica) apresentaram melhores resultados, o que pode ser confirmado na **Tabela 3**. Ao observar os resultados de concordância, foi possível identificar uma concordância moderada em considerar imagens relevantes quando os mesmos foram recuperados por meio do descritor CFM para duas análises de concordância (A x B; A x C; - **Tabela 1**). Ao observar o percentual de concordância entre os avaliadores (A x B; A x C; B x C), principalmente entre os resultados relevantes (R – R) na (**Tabela 2**), esses valores foram baixos. Porém, é possível observar maior percentual de concordância para o descritor CFM, seguido por Alldescriptors.

Os resultados deste estudo corroboram a hipótese inicial de que o descritor de forma CFM pode ser o melhor para descrever o fenômeno analisado, uma vez que apresentou

valores melhores de *precision* e *recall* para P@5, R@5, P@10 e R@10 em relação a todos os descritores avaliados, com exceção do descritor All descriptors. Estes resultados mostram que o descritor CFM pode ser mais robusto em relação à busca de imagens do que outros descritores mais simples como, por exemplo, Área/Perímetro, Circularidade e Retangularidade, que apresentaram resultados inferiores, sendo estes invariantes à escala, o que pode ser um fator para baixo desempenho destes descritores.

Os objetos analisados neste estudo são imagens de polígonos formados pelo envoltório convexo a partir de dados da posição de jogadores de futebol. Equipes de futebol apresentam magnitudes distintas de área de ocupação em contexto de jogo, por exemplo, quando os mesmos estão com e sem posse de bola (Moura et al., 2012). Neste contexto, as equipes em jogo podem apresentar diferentes formas de organização durante a partida ou até semelhantes, porém em escalas diferentes, o que torna um fenômeno complexo para a detecção de imagens, sendo necessário realizar uma transformação da imagem e detectar estas variâncias de imagens (Aslandogan e Yu, 1999). Assim, um descritor que considera a variação de escala pode ser a melhor opção para este fenômeno, como foi o caso para a CFM. Este descritor considera a representação de uma forma para diferentes escalas, ou seja, cada escala está relacionada a diferentes valores de dimensão fractal (Torres, Falcão e Costa, 2004). Outros dois descritores também foram extraídos a partir da curva, como AF e MF, porém, um único valor que descreve a forma neste estudo analisado demonstrou não ser bom, comparados a um VC.

Ao considerar o descritor CFM uma boa ferramenta para descrever o formato da organização de equipes de futebol, tornou-se possível aplicar o mesmo para a análise do formato de organização em uma partida de futebol do campeonato brasileiro, o que nos permite identificar as formas mais comuns que as equipes executaram durante a partida. Formas de organização maiores foram encontradas predominantemente (ex. **Figura 11D**, C1, da equipe 2). Em uma comparação entre curvas, as formas que são menores (ex. **Figura 11A**, C5, da equipe 1), apresentaram queda de valores de dimensão fractal antecipadamente em relação às formas maiores (ex. **Figura 11A**, C1, da equipe 1). Isto acontece devido à perda de auto-similaridade da forma em um determinado intervalo de dilatação, quando o mesmo aumenta ou diminui (Torres, Falcão e Costa, 2004).

Na literatura, é possível observar estudos que relatam o controle do espaço realizado por equipes em campo utilizando a interpretação dos valores de magnitude da área de ocupação (representada pelo envoltório convexo), aspecto central para a interpretação desses dados (Frencken et al., 2011; Moura et al., 2012; Rein e Memmert, 2016). Quando as imagens do envoltório convexo são obtidas e analisadas pelos valores da CFM, torna-se possível resgatar formas de organização que uma dada equipe realiza. No presente estudo, identificou-se que ambas as equipes apresentaram várias formas diferentes, em escalas diferentes, dependendo das diferentes situações da partida. Porém, foi possível verificar que, mesmo com uma grande variedade de forma que uma determinada equipe possa apresentar, algumas das formas são consistentes durante a partida. Por exemplo, a equipe 1 apresentou a forma rotulada “C6” durante 36% do segundo período (**Figura 11B**). Particularmente, pode-se argumentar dois importantes *feedbacks* aos jogadores e treinadores com tal análise: a) uma possibilidade de confirmar se a equipe se organizou de acordo com a estratégia previamente determinada pelo técnico e b) a possibilidade de o treinador traçar a forma de organização da equipe desejada e então buscar no banco de dados os momentos em que a equipe jogou em concordância com o desejado.

Assim, ao apresentar uma ferramenta importante como a Dimensão Fractal Multiescala para descrever a organização de equipes de futebol, foi possível extrair as formas em que as equipes de futebol realizam durante uma partida. A partir destes resultados pode-se pensar em aplicar esta ferramenta em condições importantes durante uma partida, quando uma equipe está em ataque ou defesa ou análises em situações determinantes de uma partida, como sequências ofensivas resultadas em gol, ou ainda para caracterizar as formas organizacionais nos diferentes sistemas de jogo (por exemplo, 4-4-2, 4-3-3, 3-5-2, etc.) adotados por equipes de diferentes países e competições.

É preciso compreender que esta ferramenta é apresentada para uma perspectiva de análise macro, ou seja, o formato considerado pelo envoltório convexo, e este negligencia, por exemplo, os jogadores da equipe que estão dentro do polígono. Assim, futuros estudos devem ser considerados para aprimorar esta ferramenta em análises mais detalhadas baseadas em grafos, por exemplo, de forma que considere todos os jogadores das equipes.

## **5.4 Conclusão**

Os resultados permitiram concluir que o descritor curva fractal multiescala (CFM) é o descritor mais eficaz em relação a todos os demais avaliados neste estudo, o que é importante para descrever formas em diferentes escalas. Assim, foi possível aplicar este descritor durante uma partida de futebol profissional e identificar as formas que as equipes executaram durante a partida. Esses resultados devem ser interpretados com cautela, tendo em vista que foi aplicado apenas a uma partida de futebol, para ambas as equipes. Portanto, a análise não deve ser extrapolada para todo o contexto da modalidade. Estudos futuros devem ser aplicados com um maior número de jogos para que possa caracterizar padrões de jogo de diferentes nacionalidades, bem como em diferentes contextos de jogo, para fornecer informações a treinadores sobre o desempenho da organização de suas equipes ou adversários.

## **6 Estudo Dois: Quais são as formas de organização que equipes de futebol adotam e sua relação com eventos do jogo? Identificação e análise em partidas oficiais.**

### **Resumo**

O objetivo do presente estudo foi analisar as formas de organização que equipes de futebol adotam durante partidas oficiais e em sequências ofensivas (SO) que terminam em sucesso ou fracasso para situações de ataque e defesa. As imagens dos polígonos correspondentes às formas, representadas pelo envoltório convexo, foram obtidas pelas coordenadas 2D dos jogadores para 14 partidas oficiais de futebol. O descritor de formas Dimensão Fractal Multiescala (DFM) foi calculado para as imagens e extraído os valores de Curva Fractal Multiesca (CFM), e realizada a discriminação das formas de organização das equipes entre os períodos e a frequência de formas para cada cluster quando realizaram finalizações ou desarmes. Um Índice de variabilidade (IV) de formas foi calculado pelo método *Sample Entropy* (SampEn) para as sequências ofensivas (SO) em situações de defesa e ataque. Para as SO que terminaram em desarme ou finalização, foi calculada a Distância Euclidiana entre as Curvas Fractais Multiescala (DECFM) e o IV. O Máximo valor de Distância Euclidiana (MvDE) entre as equipes durante as SO também foi extraído. Os resultados de IV de formas não apresentaram diferenças significativas em situação de defesa (SO que terminaram em desarmes realizados (SampEn = 0,23 (0,05)) comparados às SO que terminaram em finalizações sofridas (SampEn = 0,23 (0,05)),  $p = 0,7935$ ;  $ES = 0,0069$ ), como também não apresentaram diferenças em situação de ataque (SO que terminaram em finalizações realizadas (SampEn = 0,24 (0,07)) em relação às SO que terminaram em desarmes sofridos (SampEn = 0,24 (0,07)),  $p = 0,6156$ ;  $ES = 0,0132$ ). Para os resultados de variabilidade para a DECFM, foram verificados maiores valores de entropia quando as SO terminaram em desarmes (SampEn = 0,19 (0,24)) em relação às SO que terminaram em finalização (SampEn = 0,15 (0,22)), ( $p = 0,0034$ ;  $ES = 0,0770$ ) e os valores de MvDE foi maior para as SO que terminaram em finalização (SampEn = 0,57 (0,37)) em relação a quando terminaram em desarmes (SampEn = 0,49 (0,33)), ( $p < 0,001$ ;  $ES = 0,0947$ ). Pôde-se concluir que o sucesso ou o fracasso em sequências ofensivas em situação de defesa e ataque não advém da variabilidade de formas, porém, uma maior variabilidade de formas relacionadas à interação

entre as equipes para as SO que terminaram em desarmes pode indicar uma possibilidade de sucesso durante as SO e a imprevisibilidade na mudança de forma em sucesso na finalização.

**Palavras-chave:** esportes coletivos; análise de partidas; tática; dimensão fractal; variabilidade de formas.

## 6.1 Introdução

Na última década, é possível observar um aumento na compreensão de conceitos e variáveis táticas relacionadas ao futebol. Com o avanço no desenvolvimento de ferramentas computacionais aplicadas a análises de dados posicionais, em proporção ao *big data*, torna-se possível estabelecer como a tática de jogo de equipes pode ser representada e analisada durante as partidas (Rein e Memmert, 2016; Goes et al., 2020). Estas respostas vêm de inúmeros estudos que exploraram diferentes variáveis relacionadas à distância entre jogadores (Rico-González et al., 2020(a)), variáveis espaciais (Rico-González et al., 2021), bem como a posição média (centro geométrico ou centroide) de uma determinada equipe em campo (Rico-González et al., 2020(b)). Estas variáveis podem ajudar técnicos e demais profissionais das ciências dos esportes a compreender sobre área efetiva de jogo, área de ocupação, espalhamento dos jogadores da mesma equipe, *stretch index*, entre outros (Frencken et al., 2011; Moura et al., 2012; Fonseca et al., 2013; Clemente M. F. et al., 2017; Caetano et al., 2021).

Especificamente, a área de ocupação é definida como a área do menor polígono convexo formado pela posição dos jogadores, e representa o total de espaço ocupado por uma equipe em campo (Frencken et al., 2011; Moura et al., 2012). Estes valores de área de ocupação ( $m^2$ ) podem estar relacionados ao desempenho em ações determinantes no futebol. A literatura mostra que estes valores em situações defensivas e ofensivas, para ações de finalizações ou desarmes, são distintos quando ocorrem o sucesso ou o fracasso entre estas ações (Moura et al., 2012). Embora tais pesquisas tenham trazido grandes avanços em termos de análise de desempenho ao caracterizar a dinâmica das áreas de ocupação durante os jogos, bem como sua associação com eventos determinantes da partida, pode-se questionar se investigar apenas as magnitudes das áreas, sem levar em consideração suas características estruturais, limita possíveis orientações a técnicos e treinadores a respeito de como equipes de

futebol devem se organizar em campo. Nesse sentido, torna-se importante a avaliação das características de forma desses polígonos gerados.

Descritores de formas apresentados na literatura podem ajudar a compreender a complexidade da modalidade e identificar diferentes formas de organização que as equipes apresentam. Um descritor que pode ser aplicado à modalidade é a Dimensão Fractal Multiescala (DFM), descritor de forma variante à escala, que possibilita a identificação não somente de diferentes formas, mas também formas semelhantes em diferentes escalas (Torres, Falcão e Costa, 2004; Bueno et al., 2021). Esta aplicação pode ser observada em Bueno et al., (2021), que mostrou como a DFM pode ser mais eficaz que diferentes descritores de formas avaliados, aplicada para identificar ao longo de uma partida polígonos que representam a organização tática no futebol.

Além de tratar as magnitudes de área de ocupação e suas variações ao longo da partida, a estrutura das formas é imprescindível para compreensão de como atletas e equipes se organizam frente aos objetivos do jogo. Equipes de futebol são, portanto, dois sistemas complexos com auto-organização, que interagem, se adaptam e variam sua organização ao longo da partida, o que também representa um fator de desempenho (Frencken e Lemmink, 2009). Por exemplo, Moura et al., (2016) caracterizaram a interação de equipes de futebol ao avaliar a coordenação entre as séries temporais de medidas de espalhamento e reportaram que, no início de sequências ofensivas resultantes em finalizações, a equipe atacante antecipa sua organização em relação ao adversário, ou realiza um comportamento completamente oposto. Portanto, uma vez que as equipes se adaptam e variam seu comportamento ao longo do tempo para atingirem seus objetivos, pode-se questionar se a variabilidade, não apenas de magnitude, mas também da forma de organização, também não explica o sucesso ou fracasso durante jogadas de ataque e defesa.

O papel da variabilidade é estudado em diversas áreas de conhecimento, principalmente em pesquisas relacionadas a sistemas dinâmicos, biomecânica, aprendizagem e controle motor (Bartlett, Wheat e Robins, 2007). Os vários graus de liberdade envolvidos na coordenação e o controle do movimento são uma fonte potencial de variabilidade. Em biomecânica e controle motor, a variabilidade é tradicionalmente equiparada ao ruído, é considerado prejudicial ao desempenho do sistema, e normalmente é eliminado dos dados como uma fonte de erro. Na avaliação das mudanças de coordenação devido ao aprendizado, envelhecimento e doença, a presença de variabilidade é considerada um dos indicadores de

desempenho ruim (Van Emmerik, Miller e Hamill, 2014). No entanto, as ciências biológicas e físicas enfatizam aspectos benéficos e adaptativos no funcionamento do sistema. Na perspectiva tradicional, níveis superiores de habilidade, competência e saúde estão associados com diminuição da variabilidade. Porém, perspectivas atuais de sistemas dinâmicos apontam o oposto; a perda de variabilidade reflete uma diminuição da interação dos graus de liberdade, o que representa um prejuízo no funcionamento do sistema (Van Emmerik, Miller e Hamill, 2014). Para o futebol, um estudo recente mostrou a variabilidade de posicionamento de atletas ao longo de uma competição renomada da Europa e mostrou que as equipes finalistas apresentam menor variabilidade de posição média dos jogadores ao longo do campeonato (Moura et al., 2015). No entanto, a posição média em uma partida ainda não reflete de fato a capacidade da equipe em se organizar, dinamicamente, ao longo da partida.

Diante da importância do papel da investigação não apenas das magnitudes de área de ocupação de equipes de futebol, mas também da forma de organização, da interação das equipes, principalmente, da capacidade de variação dessa estrutura organizacional, levanta-se algumas questões que requerem respostas sobre aspectos táticos na modalidade: a) como equipes mudam suas formas de organização em relação ao adversário para atingir seus objetivos na partida?; b) qual é a relação entre a forma de organização e eventos determinantes da partida?; c) qual é o papel da variabilidade de forma em sequências ofensivas determinantes na partida?

O objetivo do presente estudo foi analisar as formas de organização que equipes de futebol adotam durante partidas oficiais e em janelas específicas de tempo que correspondem às sequências ofensivas (SO) que terminam em sucessos ou fracassos para situações de ataque e defesa. A hipótese inicial foi que, ao analisar a variabilidade de forma de organização tática entre as equipes em diferentes contextos da partida, como em SO para situações determinantes do jogo, haveria relações entre sucesso e maiores variações estruturais (formas) das equipes, o que pode representar a capacidade coletiva de adaptação ao meio. Além disso, a identificação de estruturas relacionadas à organização em momentos de sucesso ao realizar um desarme e uma finalização foi apresentada.

## 6.2 Materiais e métodos

### 6.2.1 Participantes

Para este estudo, a amostra constituiu na obtenção de coordenadas 2D de jogadores de futebol profissional durante partidas oficiais de campeonatos brasileiros e estaduais. Para tanto, um (1) jogo do campeonato paulista (Campeonato regional Série A-1) de 2004 foi coletado ( $n = 28$  jogadores), 9 jogos do campeonato brasileiro (Campeonato nacional Série A-1) foram coletados entre os anos de 2001/2003/2004/2008 ( $n = 249$  jogadores). Também, 4 jogos do campeonato paulista (Campeonato regional Série A-2) de 2014 ( $n = 101$  jogadores) foram coletados. As descrições das partidas analisadas neste estudo podem ser encontradas no apêndice I. O presente estudo foi aprovado pelo comitê de ética em pesquisa (parecer nº 3.047.461).

### 6.2.2 Coleta de dados

#### 6.2.2.1 *Obtenção das coordenadas 2D dos jogadores*

Foram obtidas as posições 2D dos jogadores em campo por meio do software DVídeo (Figueroa, Leite e Barros, 2006a; 2006b), onde o mesmo utiliza ferramentas baseado em vídeo para a obtenção das coordenadas 2D dos jogadores. Cada jogador, de cada time, foi numerado como  $p = 1, 2, \dots, 14$  para a equipe anfitriã e  $p = 15, 16, \dots, 28$  para a equipe adversária. Assim, as coordenadas bidimensionais dos jogadores foram definidas como  $p(Xp(t), Yp(t))$ , onde  $t$  representa cada instante de tempo (em segundos). Em seguida, um filtro digital passa-baixa Butterworth de 3ª ordem, com uma frequência de corte de 0,4 Hz foi utilizado para suavizar os dados das coordenadas 2D de todos os jogadores.

#### 6.2.2.2 *Obtenção das imagens que representam a organização das equipes em campo e o cálculo da Dimensão Fractal Multiescala (DFM)*

A partir dos dados de posição 2D dos jogadores, é possível obter as imagens que representam a organização tática das equipes ao longo do tempo ( $t$ ) de cada partida. Para representar a organização das equipes em campo, foi gerado o envoltório convexo por meio

de pontos (jogadores) com respectivas coordenadas 2D no plano cartesiano (Preparata e Shamos, 1985; Moura et al., 2012). A demonstração esquemática da obtenção das imagens da organização das equipes pode ser mais bem elucidada na seção 4.2 (**Figura 7**).

De posse das imagens da organização tática das equipes, será calculada a DFM para a descrição quantitativa destas imagens (Torres, Falcão e Costa, 2004), com extração dos valores de Curva Fractal Multiescala (CFM), para um vetor de características com amostra de 100 pontos (ver descrição detalhada em seção 4.3, com demonstração de exemplo na **Figura 8**).

### **6.2.2.3 Obtenção das ações técnicas e caracterização das sequências ofensivas (SO)**

Após a obtenção da posição 2D dos jogadores, foi realizado o registro das ações técnicas realizadas pelos jogadores durante as partidas oficiais (**Figura 3**). O software DVVideo possui interfaces para identificar qual ação (ex. finalização, desarme, etc.) um determinado jogador realizou, qual instante (frame) do jogo foi realizada esta ação, qual era a posição deste jogador em campo quando realizou a ação, e se for o caso, identificar se esta ação foi realizada de forma correta ou errada.

As informações das ações técnicas foram importantes para identificar o início e o fim das sequências ofensivas (SO) em situações de defesa e ataque. O início de uma SO foi considerado quando as equipes retomaram a posse de bola (início da partida, tiro de meta, lateral, escanteio, ou a retomada por um desarme que resultou em posse de bola) e o sucesso e fracasso para as SO em situação de defesa e ataque foram consideradas conforme descrito abaixo (Moura et al., 2012):

- **Situação de defesa:** o sucesso de uma SO, sob a perspectiva da defesa, foi considerado quando a equipe realizou um desarme e retomou a posse de bola. O fracasso foi considerado quando a SO terminou em uma finalização realizada pela equipe adversária.
- **Situação de ataque:** o sucesso, sob a perspectiva do ataque, foi considerado quando a equipe terminou a SO com a realização de uma finalização. Por outro lado, o fracasso em uma SO foi considerado quando a equipe sofre um desarme da equipe adversária.

Sequências ofensivas que foram originadas de escanteios e faltas foram excluídas devido a curta duração de tempo.

## 6.2.3 Análise dos dados

### 6.2.3.1 Aplicação do Método de Cluster

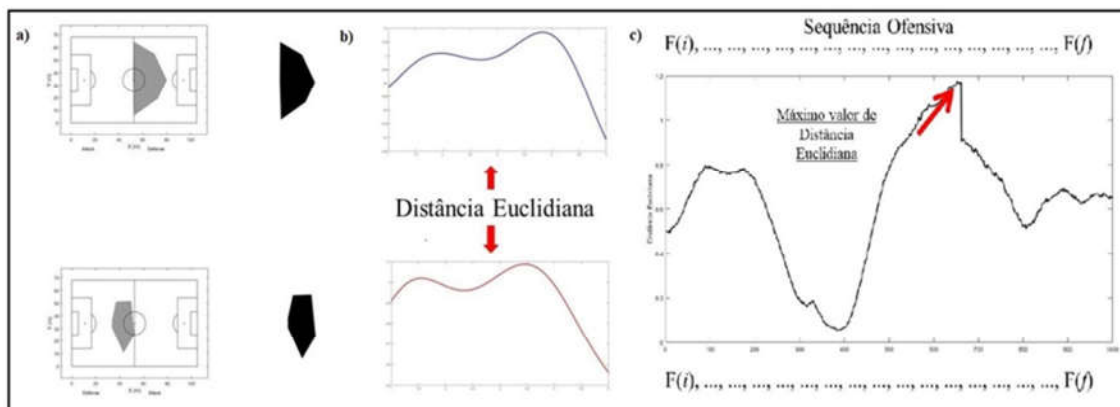
Para a discriminação das formas correspondente a organização tática que as equipes realizam durante as partidas, o método de cluster *k-means* (Kaufman e Rousseeuw, 1990) foi aplicado para identificar as formas de organização em diferentes grupos que as equipes adotam durante as partidas oficiais, para cada equipe e entre os períodos, bem como a frequência relativa (FR) de formas em situações de defesa e ataque. Este agrupamento foi realizado por meio dos dados de Distância Euclidiana (DE) dos vetores de características das CFM (Bueno et al., 2021). O método *Elbow* (Han, Kamber e Pei, 2012) foi utilizado para identificar qual o número de clusters *k* ideal a ser utilizado para a realização do método de cluster. O índice de agrupamento foi avaliado pela média do Coeficiente de Silhueta (CS) e é considerado satisfatório quando ultrapassa 0,7 (Kaufman e Rousseeuw, 1990).

### 6.2.3.2 Análise de variabilidade

De posse dos dados das SO, foi calculado o Índice de Variabilidade (IV) de formas para as SO em situações de ataque, quando realizam finalização ou sofrem desarmes e em situações de defesa, quando realizam desarmes ou sofrem finalizações. O IV foi calculado pelo método *Sample Entropy* (SampEn) para o vetor que obtém todos os valores de DFM para cada SO correspondente às situações de ataque ou defesa, com valores de  $m = 2$  e tolerância  $r$  de 0,2 (Richman e Moorman, 2000).

Também, para as SO que resultaram em finalizações e desarmes, foi calculado o IV da Distância Euclidiana entre as curvas de fractal multiescala (DECFM) das equipes, como pode ser representado no esquema na **Figura 12**. Para cada imagem de organização representada por meio do envoltório convexo (**Figura 12(a)**), é obtida a CFM (**Figura 12(b)**), e posterior calculou-se a DECFM entre as equipes durante cada SO (**Figura 12(c)**). O Máximo valor de Distância Euclidiana (MvDE) entre as curvas fractais das equipes nas SO também foi identificado.

O IV da DECFM pode indicar o grau de interação entre as equipes e representa o quanto as equipes modificam suas estruturas que antecipam uma ação determinante, como finalização ou desarme. O MvDE representa o máximo que duas equipes se diferenciam em formas durante a SO.



**Figura 12.** Esquema para a obtenção da série temporal correspondente a DECFM para as SO que resultaram em desarme ou finalização. (a), obtenção dos polígonos convexas; (b), obtenção da CFM e cálculo da DE; (c), exemplo de série temporal da DECFM e o MvDE para uma SO. (Fonte: o próprio autor).

#### 6.2.4 Análises estatísticas

Primeiramente, análises descritivas foram realizadas para demonstrar em quantas formas e quais formas possíveis as equipes de futebol se organizam durante as partidas. Para melhor visualização e compreensão dos resultados, gráficos que apresentam exemplos de forma que as equipes realizaram durante as partidas (em momentos específicos durante a partida), em qual período da partida que apresentam formas semelhantes (gráfico de barra), o percentual de formas para cada cluster (gráfico de pizza), e o percentual de ocorrência de formas em situações de ataque e defesa, foram apresentados.

Um teste de normalidade *Lilliefors* foi realizado para verificar se houve distribuição normal dos dados. Para todos os testes, não foram encontrados resultados de normalidade, portanto, testes não paramétricos foram utilizados para todas as análises seguintes.

Um teste de *Wilcoxon* foi realizado para a comparação do IV de formas para as SO em situação de defesa, ou seja, quando as equipes realizaram desarmes (sucesso) em relação quando as equipes sofreram finalizações (fracasso). Para os IV das SO em situação de ataque, foi verificado se houve diferenças quando as equipes realizaram finalizações (sucesso) em relação quando as equipes sofreram desarmes (fracasso).

O mesmo teste foi aplicado para verificar se houve diferenças dos valores de DECFM para as SO terminadas em finalizações em relação as SO terminadas em desarmes, bem como

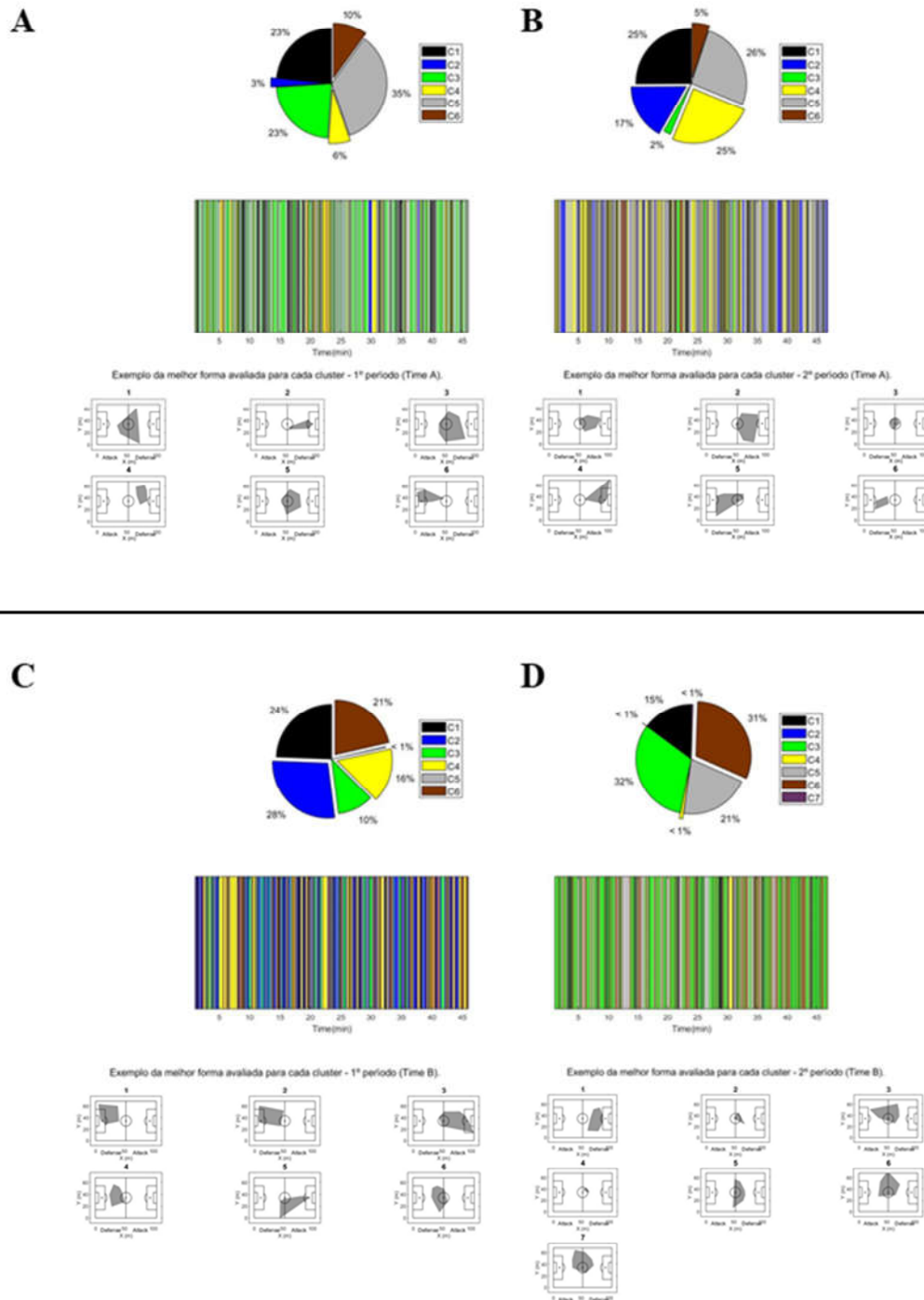
para verificar se houve diferenças na variável MvDE, entre as SO que terminam em finalizações e as SO que terminaram em desarmes.

Para todas as análises, o cálculo de *Effect Size* (ES) para análises independentes foi realizado, a considerar índices de classificação como pequeno ( $ES < 0,30$ ), médio ( $0,30 \geq ES < 0,50$ ) e grande ( $ES \geq 0,50$ ), (Cohen, 1988).

Foi adotado um valor de significância de  $p < 0,05$  e todos os resultados foram apresentados em mediana (intervalo interquartil).

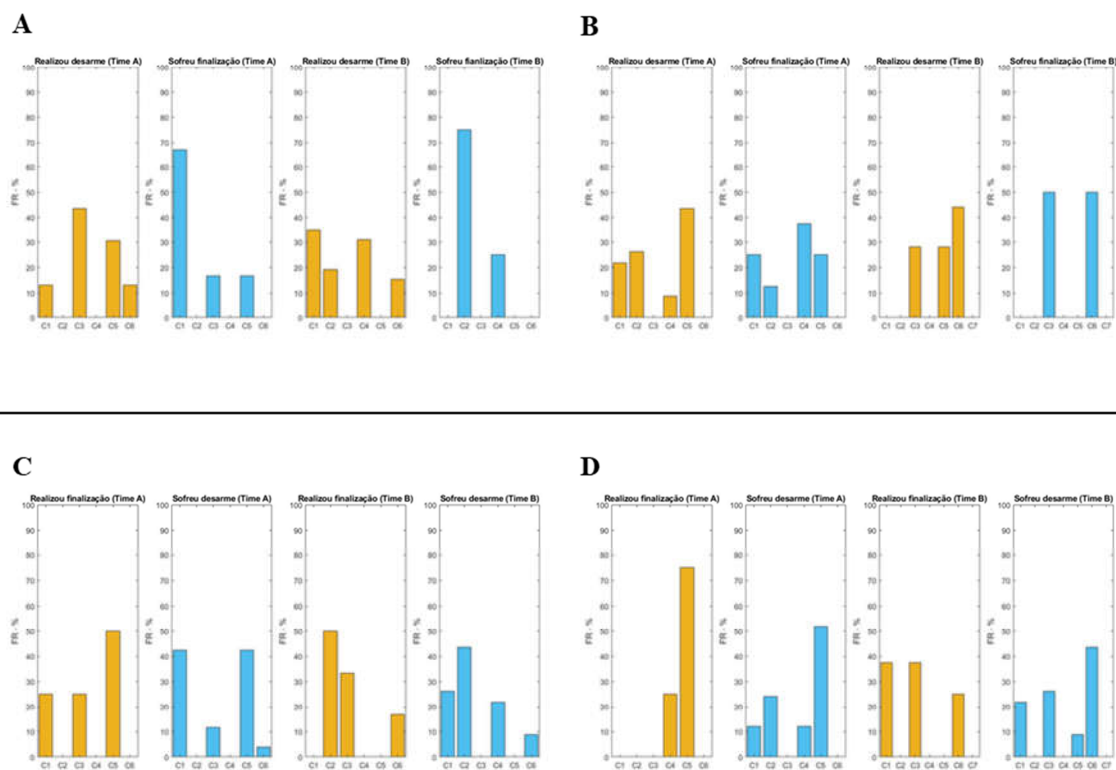
### **6.3 Resultados**

A **Figura 13** apresenta um exemplo de relatório das formas de organização para a equipe anfitriã (Equipe A) e para a equipe visitante (Equipe B). Exemplos de formas de organização que as equipes realizam durante uma partida são apresentados (1º e 2º período), bem como o percentual de formas em cada cluster e o momento de ocorrências das formas durante a partida.



**Figura 13.** Exemplos de formas de organização, percentual (*pie graphics*), e momento de ocorrências de formas (*bar graphics*) em uma partida. (A), equipe anfitriã, 1º período; (B), equipe anfitriã, 2º período; (C), equipe visitante, 1º período; e (D) equipe visitante, 2º período. (Fonte: o próprio autor).

Na **Figura 14**, pode-se observar um exemplo para a mesma partida sobre a frequência relativa (FR) de forma que as equipes atuaram, para cada cluster, em momentos específicos em situações de defesa (quando realizou desarme e sofreu finalização) e situações de ataque (quando realizou finalização e sofreu desarme).



**Figura 14.** Resultado de Frequência Relativa (FR) de formas de organização, para cada cluster, que as equipes realizam em momentos específicos de situações de ataque e defesa. (A), FR em situações de defesa, 1º período; (B), FR em situações de defesa, 2º período; (C), FR em situações de ataque, 1º período; e (D), FR em situações de ataque, 2º período. (Fonte: o próprio autor).

Os resultados da **Tabela 4** mostram os valores de IV de formas para as SO em situações de defesa e em situações de ataque. Estes resultados mostram que não houve diferenças estatísticas, em situação de defesa, para as SO que terminaram em desarmes comparados a SO que terminaram em finalizações sofridas pela equipe adversária ( $p = 0,7935$ ;  $ES = 0,0069$ ). Em situações de ataque, também não foram encontradas diferenças significativas para as SO que terminaram em finalizações em relação a desarmes sofridos pelo adversário ( $p = 0,6156$ ;  $ES = 0,0132$ ).

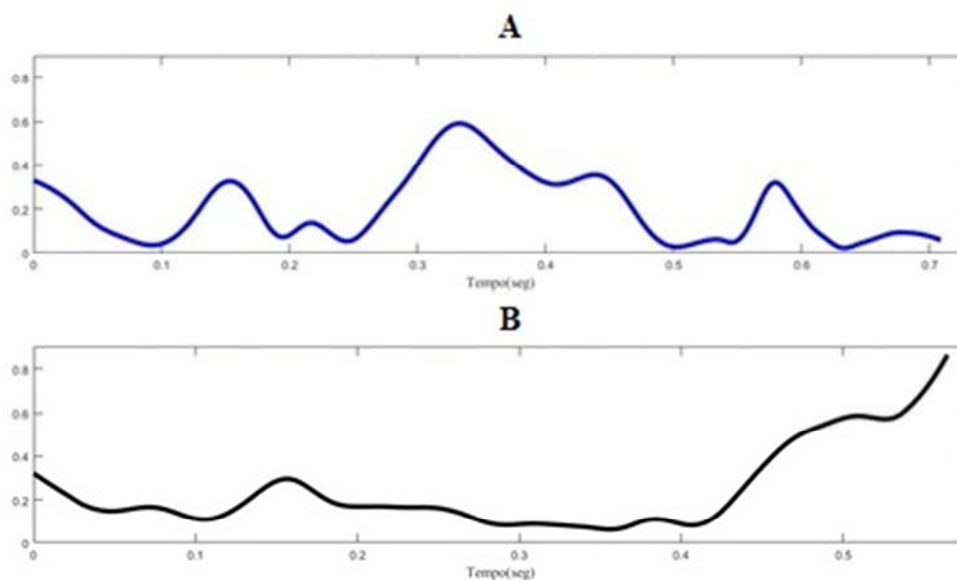
**Tabela 4.** Resultados de mediana (intervalo interquartil) para os valores de IV de formas (SampEn) em sequências ofensivas (SO) para situações de defesa ( $n = 1193$ ) e situações de ataque ( $n = 256$ ).

Variável	Situação de defesa		Situação de ataque	
	SO que terminaram em desarmes	SO que terminaram em finalizações sofridas	SO que terminaram em finalizações	SO que terminaram em desarmes sofridos
IV de formas	0,23 (0,05)	0,23 (0,05)	0,24 (0,07)	0,24 (0,07)

A **Figura 15** apresenta um exemplo de duas séries temporais que representam a DECFM entre as equipes durante uma SO que terminou em desarme (**Figura 15(a)**) e uma SO que terminou em finalização (**Figura 15(b)**). Em uma análise visual, a série temporal da SO que resultou em desarme apresenta maiores oscilações em relação à SO que resultou em finalização, sendo que maiores oscilações podem indicar uma maior variabilidade de formas empregadas pela equipe durante a SO.

Isto pode ser confirmado na **Tabela 5**, que apresenta resultados significativamente maiores de IV das DECFM para as séries temporais das SO que resultaram em desarmes comparados as séries temporais das SO que resultaram em finalizações ( $p = 0,0034$ ;  $ES = 0,0770$ ). Quando as séries temporais resultaram em finalizações, o valor de máxima distância entre as CFM foi maior em relação às séries temporais que terminaram em desarmes ( $p < 0,001$ ;  $ES = 0,0947$ ).

Também, em uma análise visual, é possível observar que o pico de distância ocorre em momentos diferentes. Por exemplo, na SO terminada em finalização, este pico ocorre no final da série temporal (**Figura 15(b)**). Este fenômeno foi possível ser observado na maioria dos dados.



**Figura 15.** Exemplo de séries temporais da distância Euclidiana das CFM entre as equipes durante uma SO que terminou em desarme (a), e uma SO que terminou em finalização (b). (Fonte: o próprio autor).

**Tabela 5.** Resultados de mediana (intervalo interquartil) dos valores de IV (SampEn) e o máximo valor extraído para os dados de Distância Euclidiana entre as Curvas Fractais Multiescala (DECFM) das equipes para as SO que terminaram em finalizações e desarmes.

Variáveis	SO que terminaram em desarmes	SO que terminaram em finalização
IV das DECFM	0,19 (0,24)*	0,15 (0,22)
MvDE	0,49 (0,33)*	0,57 (0,37)

\*: Significativamente diferente de SO que terminaram em finalizações.

Abreviações: IV = Índice de variabilidade; DECFM = Distância Euclidiana da Curva fractal Multiescala; MVDE = Máximo valor de Distância Euclidiana.

## 6.4 Discussão

O objetivo do presente estudo foi analisar as formas de organização tática que equipes de futebol adotam em janelas específicas de tempo que correspondem às sequências ofensivas (SO) que terminam em sucessos ou fracassos para situações de defesa e ataque. Para isso, a busca de formas de organização táticas das equipes de futebol, representadas pelo envoltório convexo, foi realizada pela Dimensão Fractal Multiescala (DFM), descritor de forma que possibilita descrever a complexidade, discriminar as formas de organização executadas pelas

equipes durante a partida, como também a escala de formas de organização tática, devido às informações multiescala das imagens analisadas (Bueno et al., 2021).

Especificamente para este estudo, foram abordados contextos de análises de formas de organização tática, como sequências ofensivas (SO) que resultaram em ações determinantes no futebol, ações estas como finalizações e desarmes. Uma análise descritiva foi realizada para discriminar quantas e quais formas as equipes de futebol realizaram durante as partidas (exemplos estes podem ser vistos nos resultados de formas em cada  $n$  clusters (C1, C2,..., Cn) na **Figura 13**, bem como o percentual de ocorrências destas formas durante a partida, e o momento em que estas formas foram executadas). Também a FR de formas realizadas em cada cluster, quando obtiveram o sucesso ou fracasso em situações de defesa e ataque (definição descrita no tópico 5.2.2.3) foram apresentadas (**Figura 14**). Um relatório para todos os jogos contendo estas informações foi realizado e está disponível (ver <https://figshare.com/s/27f660f9d2ddc5a98217>).

Neste estudo, foi analisada a organização tática das equipes em SO para situações de defesa e ataque. Estas situações têm desfechos bem estabelecidos pela literatura (Moura et al., 2012). Para analisar estas SO em uma perspectiva de formas de organização e não apenas em magnitude de ocupação em campo, um IV de formas foi calculado. Este IV foi calculado por meio do SampEn, medida não-linear, sendo uma de suas definições e aplicações apresentar uma métrica para prever o quão bem pode-se prever um próximo estado de um sistema. A entropia aplicada a séries temporais biológicas pode fornecer uma medida de imprevisibilidade ou irregularidade de um sistema (Stergiou, 2016).

A hipótese inicial de que a variabilidade de forma de organização tática das equipes em diferentes contextos da partida pode estar atrelado ao sucesso não foram confirmadas, uma vez que não houve diferenças de IV de formas quando as equipes obtiveram sucesso ou fracasso em situações de defesa e ataque. Ao aplicar o conceito teórico de imprevisibilidade ou irregularidade para as séries temporais dos valores de CFM correspondente às formas de organização nas SO, pode-se dizer que o sucesso ou o fracasso não é dependente de imprevisibilidade de organização de formas, individualmente.

Porém, ao se analisar IV das formas de organização entre as equipes por meio da DECFM, quando as SO obtiveram como desfecho o desarme, o IV apresentou valores maiores comparados às SO que terminaram em finalização. Estes resultados podem indicar

uma constante adaptação de forma ao longo da SO por parte da equipe defensiva (ver **Figura 15**, como exemplo), um indicativo para que consiga com maior sucesso realizar um desarme. Por outro lado, a equipe ofensiva, por estar com a posse de bola, pode proporcionar esta maior variabilidade em relação à equipe defensiva, devido aos inúmeros passes, dribles e outras ações que realizam para chegar ao objetivo de uma finalização. Estudos futuros podem ser realizados para correlacionar variáveis de ações técnicas com os valores de DECFM relacionadas às formas de organização durante SO.

Outro indicativo de sucesso, desta vez para as SO que terminaram em finalização, diz respeito ao MvDE que as equipes apresentam. Este valor (MvDE) pode indicar qual o máximo que uma equipe difere em forma de organização em relação ao adversário durante a SO. Quando as SO terminaram em finalização, foi encontrado um valor maior do IV para diferença de formas em relação quando as SO terminaram e desarme. A **Figura 15** pode ajudar a compreender melhor estes valores. Mesmo quando a magnitude do MvDE é diferente entre os desfechos, é possível observar que esta máxima diferença ocorre em momentos distintos. Quando SO que terminaram em desarmes, o MvDE ocorreu em torno do centro da série temporal. Para a maioria das SO que terminaram em finalização, ocorreu próximo ao final, visto no exemplo da **Figura 15**, o que não corrobora com os resultados de Moura et al., (2016). Os autores mostraram em análise de coordenação da distribuição de jogadores que, em SO que terminaram em finalizações, há valores de distribuição em anti-fase no início da SO, ou seja, as equipes apresentam distribuições opostas no início da SO. Neste estudo, se o comportamento visto na **Figura 15** se repetir nas outras situações, pode representar um indicativo de que a equipe que está em ataque, para realizar uma finalização, modifica a estrutura de organização próxima ao desfecho, o que acarreta em uma imprevisibilidade diante ao adversário, e assim, consegue realizar a finalização. No entanto, estes valores devem ser mais bem explorados em relação ao período em que ocorre a MvDE e a relação com o sucesso no desfecho das SO.

Técnicos e pesquisadores podem se beneficiar desta ferramenta para compreender as formas de organização tática aplicadas ao futebol. Compreender que o sucesso ou o fracasso no desfecho em situações de ataque e defesa não advêm de variabilidade de formas, porém a variabilidade de formas em relação à interação entre as equipes em SO, observada pela DECFM e a MvDE, pode representar um comportamento distinto quanto ao desfecho em desarme ou finalização. Estudos futuros devem buscar expandir outras aplicações para

predição de desempenho tático diante diferentes métodos de treinamento, análises em diferentes categorias, variabilidade durante o campeonato, diferentes nacionalidades.

## **6.5 Conclusão**

Com este estudo, por meio da DFM, foi possível identificar as formas de organização em jogos de futebol, bem como a frequência de forma que as equipes atuam em situações de defesa e ataque. Pôde-se concluir que o sucesso ou o fracasso em sequências ofensivas em situação de defesa e ataque não advém da variabilidade de formas, porém, uma maior variabilidade de formas relacionadas à interação entre as equipes para as SO que terminaram em desarmes pode indicar uma possibilidade de sucesso durante as SO e a imprevisibilidade na mudança de forma em sucesso na finalização.

## 7 Considerações finais

Como observado ao longo desta tese, diversas ferramentas biomecânicas foram e estão sendo desenvolvidas para a compreensão do futebol. Características físicas (distâncias percorridas, velocidade, sprints,...), características técnicas (análise de ações determinantes no futebol), bem como características táticas dos jogadores em campo. Esta última, em muitos estudos, busca a compreensão e caracterização da magnitude de ocupação das equipes em campo e como esta magnitude se comporta ao longo da partida ou campeonato. Esta organização muitas vezes é apresentada por meio de variáveis como espalhamento, *stretch index*, centroide, largura e comprimento, região de domínio dos jogadores e área de ocupação. Como detalhado no decorrer da tese, valores de magnitude de área de ocupação são muito bem empregadas para a descrição da organização tática de equipes de futebol, ao mostrar a ocupação espacial das equipes quando atacam ou defendem (com e sem posse de bola) e em situações determinantes da partida, como quando ocorre uma finalização ou desarme. Porém, a forma estrutural dessa organização é negligenciada e pode ser um importante fator de sucesso para a equipe durante uma partida.

Ferramentas e análise de descritores de formas foram desenvolvidas, validadas e aplicadas para a forma de organização tática de equipes de futebol em partidas oficiais. Para isso, dois estudos foram elaborados para elucidar os problemas e objetivos propostos pela tese:

O estudo um avaliou diferentes descritores de formas aplicados a imagens de polígonos que representam a organização de equipes de futebol em campo. Estes descritores de formas avaliados foram Retangularidade, Circularidade, Área/perímetro e Dimensão Fractal Multiescala. Deste último, extraiu-se valores como Área Fractal, Máximo valor Fractal, Curva Fractal Multiescala (CFM) e um último descritor com a concatenação de todos os descritores, menos a CFM, denominado *Alldescriptors*. A hipótese inicial era que a DFM apresentaria melhores resultados para descrever a forma de organização de equipes de futebol, por não ser invariante à escala. A hipótese pôde ser comprovada; entre os descritores avaliados, a CFM apresentou melhores resultados de eficácia para o fenômeno aqui estudado. Por fim, ao confirmar a hipótese, um método de cluster *k-means* foi aplicado para uma partida de futebol oficial e apresentado um exemplo de possível aplicação para saber quais são as formas de organização que equipes de futebol realizam.

No estudo dois, ao saber que a DFM é o melhor descritor, dentre os avaliados no estudo um, para descrever as formas de organização tática, a aplicação desta ferramenta pôde ser expandida e com isso, demonstrar outras possíveis aplicações para análise e interpretação da organização tática de equipes de futebol. A análise das formas de organização foi abordada em diferentes contextos da partida que resultam em ações determinantes do jogo, bem como se as formas de organização que resultam em sucesso ou fracasso durante situações de ataque e defesa. Estas análises foram feitas por meios da variabilidade de formas para sequências ofensivas a partir dos valores de CFM obtidas para as formas de organização tática. Pôde-se verificar uma relação entre variabilidade de formas e sucesso dado que as SO que resultaram em desarmes apresentaram maiores variabilidades de formas entre as equipes. Além disso, uma imprevisibilidade entre as equipes (visto pelo MvDE) também apresentaram relação com sucesso em finalizar. Estes fenômenos devem ser mais bem estudados futuramente, possivelmente na busca de se compreender quais ações técnicas produziram essas alterações súbitas de comportamento, que levaram às alterações na interação entre as equipes.

Os resultados e exemplos de formas de organização táticas apresentadas nesta tese pode não representar a forma de organização empregada hoje pelas equipes de futebol. Portanto, deve-se ter cautela ao generalizar estes resultados para o futebol contemporâneo. Técnicos e pesquisadores podem se beneficiar desta ferramenta para compreender as formas de organização tática aplicadas ao futebol e expandir outras aplicações para predição de desempenho tático diante diferentes métodos de treinamento, análises em diferentes categorias, variabilidade durante o campeonato, diferentes nacionalidades.

## 8 Referências

- Abbasi, S., Mokhtarian, F. e Kittler, J. Enhancing CSS-based shape retrieval for objects with shallow concavities. Image and Vision Computing, v.18, p.199 - 211. 2000.
- Aslandogan, Y. A. e Yu, C. T. Techniques and systems for image and video retrieval. IEEE Transactions on Knowledge and Data Engineering, v.11, n.1, p.56-63. 1999.
- Baeza-Yates, R. e Ribeiro-Neto, B. Modern Information Retrieval. 1999
- Barros, R. M. L., et al. Analysis of the distances covered by first division Brazilian soccer players obtained with an automatic tracking method. Journal of Sports Science and Medicine, v.6, p.233-242. 2007.
- Bartlett, R., et al. Analysing Team Coordination Patterns from Player Movement Trajectories in Soccer: Methodological Considerations. International Journal of Performance Analysis in Sport, v.12, n.1, p.398 - 424. 2012.
- Bartlett, R., Wheat, J. e Robins, M. Is movement variability important for sports biomechanists? Sports Biomechanics, v.6, n.2, p.224-243. 2007.
- Bueno, M. J. O., et al. How do futsal players of different categories play during official matches? A tactical approach to players' organization on the court from positional data. Plos One, v.13, n.6, p.1-13. 2018.
- Bueno, M. J. O., et al. Multiscale fractal dimension applied to tactical analysis in football: A novel approach to evaluate the shapes of team organization on the pitch. Plos One, v.16, n.9, p.1-14. 2021.
- Caetano, F. G., et al. Football player dominant region determined by a novel model based on instantaneous kinematics variables. Scientific Reports, v.11, n.1 p.1-10. 2021.
- Chaudhuri, B. B. e Sarkar, N. Texture Segmentation Using Fractal Dimension Ieee T Pattern Anal. 17: 72 - 77 p. 1995.
- Clemente M. F., et al. Measuring Collective Behaviour in Football Teams: Inspecting the impact of each half of the match on ball possession. International Journal of Performance Analysis in Sport, v.13, p.678 - 689. 2017.
- Cohen, J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, v.XX, n.1, p.37 - 46. 1960.
- Cohen, J. Statistical power analysis for the behavioral sciences. 1988
- Costa, L. F. e Cesar Jr, R. M. Shape Classification: Theori and Practice. Taylor & Francis Group. 2009

Dellal, A., et al. Comparison of physical and technical performance in European soccer match-play: FA Premier League and La Liga. European Journal of Sport Science, v.11, n.1, 2011/01/01, p.51-59. 2011.

Di Salvo, V., et al. Sprinting analysis of elite soccer players during European Champions League and UEFA Cup matches. J Sports Sci, v.28, n.14, Dec, p.1489-94. 2010.

Di Salvo, V., et al. Performance characteristics according to playing position in elite soccer. International Journal of Sports Medicine, v.28, n.3, Mar, p.222-7. 2007.

Duarte R., et al. Capturing Complex, Non-Linear Team Behaviours During Competitive Football Performance. Journal of Systems Science and Complexity, v.26, p.62 - 72. 2013.

Falcão, A. X., Costa, L. D. F. e Cunha, B. S. Multiscale skeletons by image foresting transform and its applications to neuromorphometry. Pattern Recognition, v.35, n.7, p.1571–1582. 2002.

Figueroa, P. J., Leite, N. J. e Barros, R. M. L. Background recovering in outdoor image sequences: An example of soccer players segmentation. Image and Vision Computing, v.24, n.4, p.363-374. 2006a.

\_\_\_\_\_. Tracking soccer players aiming their kinematical motion analysis. Computer Vision and Image Understanding, v.101, n.2, p.122-135. 2006b.

Fisher, Y. Fractal image compression : theory and application. 1995 (New York: Springer-Verlag.)

Fonseca, S., et al. Measuring spatial interaction behavior in team sports using superimposed Voronoi diagrams. International Journal of Performance Analysis in Sport, v.13, n.1, p.179–189. 2013.

Frencken, W., et al. Oscillations of centroid position and surface area of soccer teams in small-sided games. European Journal of Sport Science, v.11, n.4, p.215-223. 2011.

Frencken, W. G. P. e Lemmink, K. a. P. M. Team kinematics of small-sided soccer games. In: T. Reilly e F. Korkusuz (Ed.). Science and Football VI. New York: Routledge, 2009, p.161-166

Goes, F. R., et al. Unlocking the potential of big data to support tactical performance analysis in professional soccer: A systematic review. European Journal of Sport Science, v.21, n.4, p.481-496. 2020.

Gonçalves, M. A., et al. The effectiveness of automatically structured queries in digital libraries. ACM/IEEE Joint Conference on Digital Libraries, p.98 - 107. 2004.

Han, J., Kamber, M. e Pei, J. Data Mining: Concepts and Techniques: Elsevier, v.3rd ed. 2012. 443 - 495 p.

Kaufman, L. e Rousseeuw, P. J. Finding groups in data : an introduction to cluster analysis. New York: Wiley. 1990. xiv, 342 p. p. (Wiley series in probability and mathematical statistics. Applied probability and statistics,)

Landis, J. R. e Koch, G. G. The measurement of observer agreement for categorical data. Biometrics, v.33, n.1, p.159 - 174. 1977.

Locaric, S. A survey of shape analysis techniques. Pattern Recogn, v.31, n.8, p.983-1001. 1998.

Low, B., et al. A Systematic Review of Collective Tactical Behaviours in Football Using Positional Data. Sports Medicine, v.50, n.2, p.1-43. 2020.

Mandelbrot, B. B. The fractal geometry of nature. San Francisco: W.H. Freeman. 1982

Memmert, D., et al. A tactical comparison of the 4-2-3-1 and 3-5-2 formation in soccer: A theory-oriented, experimental approach based on positional data in an 11 vs. 11 game set-up. PLOS ONE, v.14, n.1, p.1-12. 2019.

Misuta, M. S. Analysis of the automatic tracking of players in collective sports. (Phd thesis). University of Campinas, Campinas, 2009. 130 p.

Misuta, M. S., et al. Representation and analysis of soccer players' trajectories. In: XXth Congress of the International Society of Biomechanics. Cleveland, USA, 2005. 415.

Mitchell, S. A. Approaches to Teaching Games: Improving invasion game performance. The Journal of Physical Education, Recreation and Dance, v.2, n.2, p.30-33. 1996.

Mokhtarian, F., Abbasi, S. e Kittler, J. Efficient and robust retrieval by shape content through curvature scale space. Proceedings of the First International Workshop on Image Database and Multimedia Search, p.35 - 42. 1996.

Moura, F. A., et al. Quantitative analysis of Brazilian football players' organisation on the pitch. Sports Biomechanics, v.11, n.1, p.85-96. 2012.

Moura, F. A., et al. Analysis of Soccer Players' Positional Variability During the 2012 UEFA European Championship: A Case Study. Journal of Human Kinetics, v.47, p.225-236. 2015.

Moura, F. A., et al. Coordination analysis of players' distribution in football using cross-correlation and vector coding techniques. Journal of Sports Sciences, v.34, n.34, p.2223-2232. 2016.

Normant, F. e Tricot, C. Method for evaluating the fractal dimension of curves using convex hulls. Physical Review A, v.43, n.12, p.6518-6525. 1991.

Penatti, O. a. B., Valle, E. e Torres, R. S. Comparative study of global color and texture descriptors for web image retrieval. Journal of Visual Communication and Image Representation, v.23, p.359 - 380. 2012.

Plotze, R. O., et al. Leaf shape analysis using the multiscale minkowski fractal dimension, a new morphometric method: a study with passiflora (passifloraceae). Canadian Journal of Botany, v.83, p. 287–301. 2005.

Preparata, F. P. e Shamos, M. I. Computational geometry : an introduction. New York: Springer-Verlag. 1985. xii, 390 p. p. (Texts and monographs in computer science)

Rein, R. e Memmert, D. Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. SpringerPlus, v.15, n.1410 p.1-13. 2016.

Richman, J. S. e Moorman, J. R. Physiological time-series analysis using approximate entropy and sample entropy. AJP – Heart and Circulatory Physiology, v.278, n.6, p. 2039–2049. 2000.

Rico-González, M., et al. Identification, Computational Examination, Critical Assessment and Future Considerations of Distance Variables to Assess Collective Tactical Behaviour in Team Invasion Sports by Positional Data: A Systematic Review. International Journal of Environmental Research, v.17, p.1-14. 2020(a).

Rico-González, M., et al. Identification, Computational Examination, Critical Assessment and Future Considerations of Spatial Tactical Variables to Assess the Use of Space in Team Sports by Positional Data: A Systematic Review. Journal of Human Kinetics, v.77, p.205-221. 2021.

Rico-González, M., et al. Origin and Modifications of the Geometrical Centre to Assess Team Behaviour in Team Sports: A Systematic Review. Revista Internacional de Ciencias del Deporte, v.16, p.318–329. 2020(b).

Sarkar, N. e Chaudhuri, B. An efficient differential box-counting approach to compute fractal dimension of image. IEEE Transactions on Systems, Man and Cybernetics, v.24, p.115 - 120. 1994.

Sarmiento, H., et al. What Performance Analysts Need to Know About Research Trends in Association Football (2012–2016): A Systematic Review. Sports Medicine, v.48, n.4, p.1-38. 2017.

Stergiou, N. Nonlinear Analysis of Human Movement Variability. CRC Press. 2016. 405 p.

Szwarcfiter, J. L. Grafos e algoritmos computacionais. Rio de Janeiro: Campus. 1984

Torres, R. S. e Falcão, A. X. Content-based image retrieval: Theory and applications. Revista de Informática Teórica e Aplicada, v.13, n.2, p.161 - 185. 2006.

Torres, R. S., Falcão, A. X. e Costa, L. F. A graph-based approach for multiscale shape analysis. Pattern Recognition, v.37, p.1163 - 1174. 2004.

Van Emmerik, R. E. A., Miller, R. H. e Hamill, J. Dynamical Systems Analysis of Coordination. In: D. G. E. Robertson, G. E. Caldwell, J. Hamill, G. Kamen e S. N. Whittlesey (Ed.). Research methods in biomechanics. Champaign, IL: Human Kinetics, 2014, p.291-315

Vilar, L., et al. Science of Winning Soccer: Emergent Pattern-Forming Dynamics in association football. Journal of Systems Science and Complexity, v.26, n.1, p.1 - 13. 2013.

Zhang, D. S. e Lu, G. J. Review of shape representation and description techniques. Pattern Recognition, v.37, p.1 - 19. 2004.

## **Apêndices**

**Apêndice I.** Descrição geral dos jogos coletados

<b>Jogos</b>	<b>Times</b>	<b>Ano de ocorrência</b>	<b>Campeonatos</b>	<b>Dimensões do campo (comprimento x largura) em metros</b>	<b>Nº de câmeras utilizadas para coleta de dados</b>	<b>Frequência de aquisição (hz)</b>	<b>Nº de Jogadores rastreados</b>
<b>Jogo 01</b>	time 01 x time 02	2001	Brasileiro – 1ª divisão	111 x 80	4	7,5	28
<b>Jogo 02</b>	time 03 x time 04	2003	Brasileiro – 1ª divisão	104 x 71	4	7,5	28
<b>Jogo 03</b>	time 05 x time 06	2003	Brasileiro – 1ª divisão	107 x 73	4	7,5	28
<b>Jogo 04</b>	time 07 x time 08	2004	Brasileiro – 1ª divisão	107 x 73	4	7,5	26
<b>Jogo 05</b>	time 09 x time 10	2004	Brasileiro – 1ª divisão	107 x 73	6	7,5	28
<b>Jogo 06</b>	time 11 x time 12	2008	Paulista – 1º divisão	106 x 72	6	7,5	28
<b>Jogo 07</b>	time 13 x time 14	2008	Brasileiro – 1ª divisão	105 x 68	6	30	27
<b>Jogo 08</b>	time 15 x time 16	2008	Brasileiro – 1ª divisão	105 x 68	6	30	28
<b>Jogo 09</b>	time 17 x time 18	2008	Brasileiro – 1ª divisão	105 x 68	6	30	28
<b>Jogo 10</b>	time 19 x time 20	2008	Brasileiro – 1ª divisão	105 x 68	6	30	28
<b>Jogo 11</b>	time 21 x time 22	2014	Paulista – 2º divisão	105 x 68	4	30	27
<b>Jogo 12</b>	time 23 x time 24	2014	Paulista – 2º divisão	105 x 68	4	30	24
<b>Jogo 13</b>	time 25 x time 26	2014	Paulista – 2º divisão	105 x 68	4	30	26
<b>Jogo 14</b>	time 27 x time 28	2014	Paulista – 2º divisão	105 x 68	4	30	24

## Apêndice II.

### Termo de Consentimento Livre e Esclarecido

#### **“Análise comparativa de descritores de formas para análises de polígonos que simulem virtualmente as organizações táticas de equipes de futebol”**

Prezado(a) Senhor(a):

Gostaríamos de convidá-lo (a) para participar da pesquisa **“análise comparativa de descritores de formas para análises de polígonos que simulem virtualmente as organizações táticas de equipes de futebol”**, a ser realizada no **“Laboratório de Biomecânica Aplicada - LBA”**. O objetivo da pesquisa é **“Realizar uma análise comparativa de descritores de formas para análise de polígonos que simulem as organizações táticas de equipes de futebol por meio de usuários reais”**.

Sua participação é muito importante e ela se daria da seguinte forma: Ao chegar ao laboratório, você será direcionado a uma sala individual, onde será explicado o objetivo do estudo e os procedimentos para a realização do teste, bem como é o funcionamento da *interface* que irá utilizar. De frente a um computador reservado só para o teste, você irá visualizar 7 imagens de polígonos, sendo uma imagem modelo e 6 imagens para avaliação, onde irá identificar quais entre as 6 imagens são mais relevantes em relação a uma imagem modelo. Após identificar as imagens que considera relevante, o mesmo clica na imagem e esta será rotulada, sendo esta identificada por você. Posterior, passa-se a analisar outras 6 imagens que possam ser relevantes para a mesma imagem modelo. No total, você será instruído a avaliar 84 imagens para três distintas imagens modelo, totalizando 252 imagens analisadas. Não será necessária a extração de fotos e vídeos durante a coleta de dados.

Esclarecemos que sua participação é totalmente voluntária, podendo você: recusar-se a participar, ou mesmo desistir a qualquer momento, sem que isto acarrete qualquer ônus ou prejuízo à sua pessoa. Esclarecemos, também, que suas informações serão utilizadas somente para os fins desta pesquisa e serão tratadas com o mais absoluto sigilo e confidencialidade, de modo a preservar a sua identidade.

Esclarecemos ainda, que você não pagará e nem será remunerado(a) por sua participação. Garantimos, no entanto, que todas as despesas decorrentes da pesquisa serão ressarcidas, quando devidas e decorrentes especificamente de sua participação. Os benefícios esperados deste estudo permite que a avaliação e validação desta ferramenta possam ajudar no desenvolvimento de ferramentas capazes de fornecer informações sobre padrões do formato de organização tática de equipes de futebol e assim fornecer subsídios para técnicos melhorar seus treinamentos. Quanto aos riscos, são mínimos, podem somente acarretar algum cansaço ou desconforto visual, devido a ficar um período a observar imagens em um computador. Caso seja necessário atendimento médico, você será acompanhado pelo pesquisador principal até o atendimento.

Caso você tenha dúvidas ou necessite de maiores esclarecimentos poderá nos contatar Murilo José de Oliveira Bueno, Rua França, 80, Jd. Adriana I, Londrina – Pr, celular, (043) 996647262, e-mail: murilo\_bu@hotmail.com, ou procurar o Comitê de Ética em Pesquisa Envolvendo Seres Humanos da Universidade Estadual de Londrina, situado junto ao LABESC – Laboratório Escola, no Campus Universitário, telefone 3371-5455, e-mail: [cep268@uel.br](mailto:cep268@uel.br).

Este termo deverá ser preenchido em duas vias de igual teor, sendo uma delas devidamente preenchida, assinada e entregue à você.

Londrina, \_\_\_ de \_\_\_\_\_ de 201\_.

### **Pesquisador Responsável**

RG: \_\_\_\_\_

\_\_\_\_\_ (NOME POR EXTENSO DO PARTICIPANTE DA PESQUISA), tendo sido devidamente esclarecido sobre os procedimentos da pesquisa, concordo em participar **voluntariamente** da pesquisa descrita acima.

Assinatura (ou impressão dactiloscópica): \_\_\_\_\_

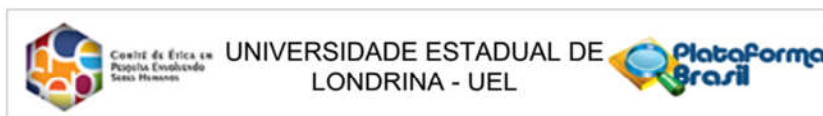
Data: \_\_\_\_\_

Obs.: Caso o participante da pesquisa seja menor de idade, o texto deve estar voltado para os pais e deve ser incluído ainda, campo para assinatura do menor e do responsável.

## **Anexos**

## Anexo I.

# Parecer do comitê de ética e pesquisa da Universidade Estadual de Londrina – UEL.



### PARECER CONSUBSTANCIADO DO CEP

#### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** Análise comparativa de descritores de formas para análises de polígonos que simulem virtualmente as organizações táticas de equipes de futebol

**Pesquisador:** Murilo José de Oliveira Murilo Bueno

**Área Temática:**

**Versão:** 2

**CAAE:** 02069518.1.0000.5231

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

**Patrocinador Principal:** Financiamento Próprio

#### DADOS DO PARECER

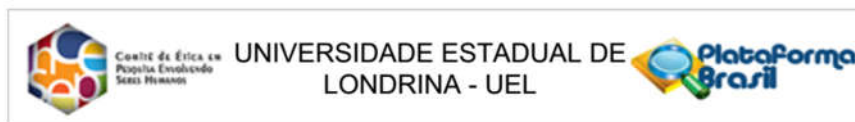
**Número do Parecer:** 3.047.461

#### Apresentação do Projeto:

Resumo:

Segundo o pesquisador, o presente artigo tem como objetivo realizar uma análise comparativa de descritores de formas para análise de polígonos que simulem as organizações táticas de equipes de futebol, especificamente, avaliar a eficácia destes descritores de forma por meio de participantes reais. Para tanto, serão convidados 30 participantes de ambos os sexos para a aplicação do teste de avaliação de relevância de imagens em relação a diferentes imagens modelos que possam simular a organização de equipes de futebol em campo. As imagens (polígonos) analisadas serão geradas por meio do método de envoltório convexo. Um método de recuperação de imagens baseadas em conteúdo será realizada para seleção das imagens a serem avaliadas realizando as seguintes etapas: Cálculo dos descritores de formas (obtenção dos vetores de características (VC) para as imagens analisadas e imagens modelo); comparação dos VC da imagem modelo e imagens analisadas (cálculo euclidiano); ranqueamento dos resultados do calculo euclidiano do menor para o maior valor; e posterior análise por meio dos participantes. Os participantes, em uma interface criada especificamente para o estudo, irão analisar 84 imagens, para 3 diferentes modelos de imagens apresentadas, totalizando 252 imagens analisadas. Assim que o participante terminar a avaliação uma matriz de dados será gerada e salva com as informações de quais imagens o participante considera relevante. Para as análises dos dados, uma

**Endereço:** LABESC - Sala 14  
**Bairro:** Campus Universitário  
**UF:** PR **Município:** LONDRINA **CEP:** 86.057-970  
**Telefone:** (43)3371-5455 **E-mail:** cep268@uel.br



Continuação do Parecer: 3.047.461

análise de eficácia dos descritores de formas será realizado para verificar qual o melhor destes descritores foi mais eficaz para recuperar imagens relevantes.

**Objetivo da Pesquisa:**

Objetivos:

Segundo o pesquisador, o objetivo principal é apresentar o uso de abordagens de descritores de formas, para caracterizar e descrever imagens virtuais que possam estar associadas a organização de equipes de futebol. • Avaliar a eficácia destes descritores de forma por meio de usuários reais que irão analisar as imagens de polígonos associadas a possíveis organizações de equipes de futebol.

**Avaliação dos Riscos e Benefícios:**

Riscos:

Segundo o pesquisador, os riscos para os participantes são mínimos, podem somente acarretar algum cansaço ou desconforto visual, devido a ficar um período observando imagens em um computador. Caso seja necessário atendimento médico, o participante será acompanhado pelo pesquisador principal até o atendimento.

**Comentários e Considerações sobre a Pesquisa:**

Benefícios:

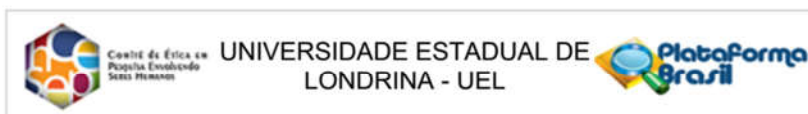
Segundo o pesquisador, a avaliação e validação desta ferramenta pode ajudar no desenvolvimento de ferramentas capazes de fornecer informações sobre padrões do formato de organização tática de equipes de futebol e assim fornecer subsídios para técnicos melhorar seus treinamentos.

**Considerações sobre os Termos de apresentação obrigatória:**

Documentos obrigatórios:

01. projeto na versão brochura
02. Termo de consentimento livre e esclarecido preenchido corretamente
03. Folha de rosto assinado pelo Vice coordenador do Programa de Pós graduação em Educação Física

Endereço: LABESC - Sala 14  
 Bairro: Campus Universitário CEP: 86.057-970  
 UF: PR Município: LONDRINA  
 Telefone: (43)3371-5455 E-mail: cep268@uel.br



Continuação do Parecer: 3.047.461

**Recomendações:**

Sem recomendações.

**Conclusões ou Pendências e Lista de Inadequações:**

O pesquisador atendeu as pendências apresentadas no parecer anterior. Recomenda-se aprovação.

**Considerações Finais a critério do CEP:**

Não existem pendências mais a serem apresentadas.

**Este parecer foi elaborado baseado nos documentos abaixo relacionados:**

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1246546.pdf	22/11/2018 20:10:02		Aceito
Projeto Detalhado / Brochura Investigador	projetodepesquisa.docx	22/11/2018 20:07:14	Murilo José de Oliveira Murilo Bueno	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	termo_consentimento.doc	29/10/2018 18:34:43	Murilo José de Oliveira Murilo Bueno	Aceito
Folha de Rosto	folha_de_rosto.pdf	29/10/2018 18:27:43	Murilo José de Oliveira Murilo Bueno	Aceito

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

LONDRINA, 29 de Novembro de 2018

Assinado por:  
Clisia M. Carreira  
(Coordenador(a))

Endereço: LABESC - Sala 14  
Bairro: Campus Universitário  
UF: PR Município: LONDRINA  
Telefone: (43)3371-5455  
CEP: 86.057-970  
E-mail: cep268@uel.br

## Anexo II.

## PLOS ONE

## RESEARCH ARTICLE

# Multiscale fractal dimension applied to tactical analysis in football: A novel approach to evaluate the shapes of team organization on the pitch

Murilo José de Oliveira Bueno<sup>1\*</sup>, Maisa Silva<sup>2‡</sup>, Sergio Augusto Cunha<sup>3‡</sup>, Ricardo da Silva Torres<sup>4‡</sup>, Felipe Arruda Moura<sup>1‡</sup>

**1** Sport Sciences Department, State University of Londrina, Londrina, Brazil, **2** Institute of Computing, University of Campinas, Campinas, Brazil, **3** College of Physical Education, University of Campinas, Campinas, Brazil, **4** Department of ICT and Natural Sciences, NTNU—Norwegian University of Science and Technology, Ålesund, Norway

✉ These authors contributed equally to this work.

‡ MS, SAC and RST also contributed equally to this work.

\* [muriljobueno@gmail.com](mailto:muriljobueno@gmail.com)



## OPEN ACCESS

**Citation:** Bueno MJdO, Silva M, Cunha SA, Torres RdS, Moura FA (2021) Multiscale fractal dimension applied to tactical analysis in football: A novel approach to evaluate the shapes of team organization on the pitch. PLoS ONE 16(9): e0256771. <https://doi.org/10.1371/journal.pone.0256771>

**Editor:** Filipe Manuel Clemente, Instituto Politécnico de Viana do Castelo, PORTUGAL

**Received:** March 19, 2021

**Accepted:** August 14, 2021

**Published:** September 1, 2021

**Copyright:** © 2021 Bueno et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data can be found at the following repository: <https://doi.org/10.6084/m9.figshare.14249456.v1>.

**Funding:** CNPq (protocol number 446548/2014-6); Fundação Araucária (protocol number 25372/2013); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001; São Paulo Research Foundation (FAPESP) grants# 2016/50250-1 -2017/20945-0 -2018/19007-9; Superintendência de Ciência,

## Abstract

The aim of this study was to evaluate different shape descriptors applied to images of polygons that represent the organization of football teams on the pitch. The effectiveness of different shape descriptors (area/perimeter, fractal area, circularity, maximum fractal, rectangularity, multiscale fractal curve—MFC), and the concatenation of all shape descriptors (except MFC), denominated AllDescriptors (AllD)) was evaluated and applied to polygons corresponding to the shapes represented by the convex hull obtained from players' 2D coordinates. A content-based image retrieval system (CBIR) was applied for 25 users (mean age of  $31.9 \pm 8.4$  years) to evaluate the relevant images. Measures of effectiveness were used to evaluate the shape descriptors (P@n and R@n). The MFD (P@5,  $0.46 \pm 0.37$  and P@10,  $0.40 \pm 0.31$ ,  $p < 0.001$ ; R@5,  $0.14 \pm 0.13$  and R@10,  $0.24 \pm 0.19$ ,  $p < 0.001$ ) and AllD (P@5 =  $0.43 \pm 0.36$  and P@10 =  $0.39 \pm 0.32$ ,  $p < 0.001$ ; R@5 =  $0.13 \pm 0.11$  and R@10 =  $0.24 \pm 0.20$ ,  $p < 0.001$ ) descriptors presented higher values of effectiveness. As a practical demonstration, the best evaluated shape descriptor (MFC) was applied for tactical analysis of an official match. K-means clustering technique was applied, and different shapes of organization could be identified throughout the match. The MFC was the most effective shape descriptor in relation to all others, making it possible to apply this descriptor in the analysis of professional football matches.

## Introduction

With the possibility of obtaining data on the athletes' position as function of time, analysing the tactical formation of players of teammates and opponents, as well as patterns of play, has been the aim of recent research on football [1–5]. Several computational methods have been