

DOI: 10.13718/j.cnki.xdzk.2025.10.018

李婷文, 张建华. 基于视频图像阈值分割的网球运动员眼球运动轨迹捕捉优化研究 [J]. 西南大学学报(自然科学版), 2025, 47(10): 208-220.

# 基于视频图像阈值分割的网球运动员 眼球运动轨迹捕捉优化研究

李婷文<sup>1,2</sup>, 张建华<sup>3</sup>

1. 闽南师范大学 体育学院, 福建 漳州 363000; 2. 湖北休闲体育发展研究中心, 武汉 430062;
3. 西北师范大学 体育学院, 兰州 730070

**摘要:** 基于视频图像阈值分割的优化方法, 通过提升眼球运动轨迹捕捉的准确性和实时性来助力运动员训练。以闽南师范大学网球运动员为研究对象, 采用 GoPro HERO12 Black 相机采集视频数据, 利用 Tsallis 相对熵进行多阈值分割获取前景图像, 并构建融合卷积神经网络与注意力机制的深度学习模型提取轨迹特征。模型引入正则项抑制过拟合, 优化特征输出。结果表明: 该方法能有效分割运动员视频前景, 生成与真实轨迹高度吻合的清晰运动轨迹, 验证了其阈值分割优势; 优化后模型损失值随训练轮次收敛至 0, 在训练效率和捕捉精度上表现优异, 具备稳定的轨迹输出能力。

**关键词:** 视频图像; 阈值分割; 网球运动员; 眼球运动轨迹;  
捕捉优化

中图分类号: TP391

文献标识码: A

开放科学(资源服务)标识码(OSID):



文章编号: 1673-9868(2025)10-0208-13

## Optimization of Eye Movement Trajectory Tracking for Tennis Players Based on Video Image Threshold Segmentation

LI Tingwen<sup>1,2</sup>, ZHANG Jianhua<sup>3</sup>

1. School of Physical Education, Minnan Normal University, Zhangzhou Fujian 363000, China;
2. Hubei Leisure Sports Development Research Center, Wuhan 430062, China;
3. School of Physical Education, Northwest Normal University, Lanzhou 730070, China

**Abstract:** Accurate eye movement trajectory tracking of tennis players remains challenging due to rapid eye

收稿日期: 2025-01-16

基金项目: 国家社会科学基金项目(20BTY090); 教育部人文社会科学基金项目(24YJCZH145); 闽南师范大学研究生教育教学改革研究项目(YJG202512); 湖北休闲体育发展研究中心开放基金项目(2022A008)。

作者简介: 李婷文, 博士, 讲师, 硕士研究生导师, 主要从事运动技术评估与优化研究。

通信作者: 张建华, 二级教授, 博士研究生导师。

motion and concentrated contribution of fixation points. The existing segmentation algorithms are difficult to accurately identify the trajectory due to local grayscale similarity, which affects the precision. To address this, we propose an optimized video image threshold segmentation method to enhance tracking accuracy and real-time performance for athlete training. Using GoPro HERO12 Black cameras, we captured video data from professional tennis players at Minnan Normal University. The foreground images were obtained via Tsallis relative entropy-based multi-threshold segmentation and fed into a deep learning model combining convolutional neural networks and attention mechanisms for feature extraction and optimization. A regularization term was introduced to mitigate overfitting and refine feature output. Experimental results demonstrate that our method effectively segmented the foreground images, producing clear and accurate eye movement trajectories that closely aligned with the ground truth, validating its superior threshold segmentation capability. The optimized model exhibits stable convergence, with loss values approaching zero with the training rounds, ensuring high efficiency and precision in trajectory tracking.

**Key words:** video image; threshold segmentation; tennis players; eye movement trajectory; tracking optimization

眼球运动主要分为平滑追踪(追随)、快速眼跳(扫视)和注视 3 种类型<sup>[1]</sup>。在网球这项高速、高对抗性的运动中,运动员需要迅速从对手的站位、引拍、挥拍击球等一系列复杂动作中准确追踪网球的运动轨迹并做出反应,以便更好地控制球的方向和速度,实现对来球方向、速度、落点及弹跳等要素的精准预判<sup>[2-3]</sup>,从而在比赛中取得优势。这一过程高度依赖于运动员的视觉搜索效率和眼球运动轨迹的精确性<sup>[4]</sup>。然而,传统的人工观察或机械记录法不仅耗时费力,且难以实时、准确地捕捉运动员在比赛中的眼球运动细节。

为实现高效、准确的眼球运动轨迹捕捉,诸多学者展开了分析。例如,研究多人情景下眼球视觉搜索特征的捕捉方法,通过采集目标人物眼睛图像,使用通光孔—角膜反射法计算眼球角膜外表面的普尔钦斑,并依据亮光点与瞳孔中心的向量关系估算眼球运动轨迹<sup>[5]</sup>。但在多人情况下光线环境存在差异,该方法的眼球运动轨迹捕捉精度欠佳。在眼动追踪技术研究方面,前期学者通过采集引航作业人员眼动数据,计算眼动范围、扫描时间和平均注视时间等参数,获得眼球运动轨迹捕捉结果<sup>[6]</sup>。然而,追踪过程中眼球存在微颤和漂移现象,使得捕捉到的轨迹不平滑不连续,影响准确性。为了探讨以用户视线为线索的眼动追踪方法,研究人员通过采集用户眼动数据,使用逻辑回归模型进行预测,得到眼球运动轨迹结果<sup>[7]</sup>。但是模型在处理速度上存在延迟,影响追踪效果。同时,也有研究者尝试基于深度学习的标记点序列预测补充方法,利用双向长短期记忆网络和时间反演对称性,通过组合损失函数约束人体运动<sup>[8]</sup>。该方法的缺点是人体运动的复杂性和多样性导致数据质量参差不齐,低质量数据增加了处理难度。

由于网球比赛视频环境复杂,背景干扰多和运动员移动速度快,深度学习模型难以识别网球运动员的眼球区域,进而无法捕捉眼球的运动轨迹。而视频图像阈值分割方法能够有效分割出目标区域和背景区域<sup>[9-10]</sup>。通过眼球区域和背景区域在灰度级上存在的差异,将目标区域从背景区域分离。使得后续针对目标区域的捕捉操作可以更加直接、高效地进行,避免了在整个图像中盲目搜索目标区域所带来的时间和资源浪费。因此,使用视频图像分割技术处理网球运动员的视频图像,并将其作为深度学习模型的输入,可以有效提高眼球运动轨迹捕捉的准确性和实时性。通过分析眼球运动轨迹捕捉结果,帮助运动员赛后复盘,优化训练计划,提升竞技水平。

# 1 研究对象与方法

## 1.1 实验名称

视频图像阈值分割下网球运动员眼球运动轨迹捕捉。

## 1.2 时间及地点

网球运动员眼球运动轨迹捕捉实验于 2024 年 9 月至 12 月在漳州市奥体中心网球场完成。

## 1.3 实验对象及实验环境

以闽南师范大学体育学院网球运动员作为实验对象。在运动员进行网球训练和比赛过程中, 使用 GoPro HERO12 Black 运动相机采集网球运动员视频图像。采集环境如图 1 所示。

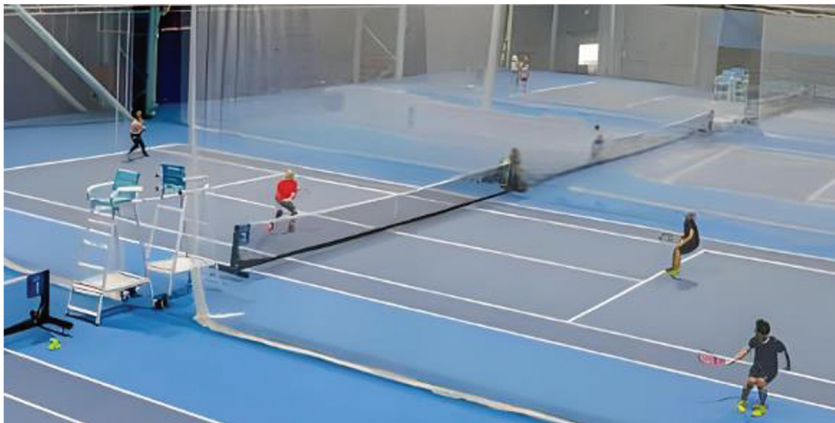


图 1 采集环境

网球运动员视频图像采集过程中所使用的相机参数如表 1 所示。

表 1 相机参数

参数类别	参数数值
传感器类型	CMOS
传感器尺寸/英寸	1.9 <sup>-1</sup>
有效像素/(百万像素 MP)	27.6(动态), 27.13(静态)
镜头光圈	F2.5
视频帧率/ fps	240/200/120/100/60/50/30/25/24
防抖性能	HyperSmooth 超强防抖 6.0
HDR	支持 HDR 视频+照片
照片模式	标准照片、连拍照片、夜景照片
视频模式	视频压缩标准为 H.265(HEVC)
照片分辨率/MP	27.13
视频最大码率/Mbps	120
视野/(°)	156(8:7 纵横比)
等效焦距/mm	12~39

由此可得如图 2 所示的网球运动员视频图像。

如图 2 所示, 借助 GoPro HERO12 Black 高性能运动相机, 采集到了网球运动员的视频图像。这些图像具备高清画质, 能流畅记录动态画面, 并拥有宽广视角, 可精准捕捉运动员的细微动作及周围环境

细节。通过先进的色彩校正和曝光控制技术,画面色彩自然真实,即便在复杂光线条件下也能保持色彩平衡和细节层次。此外,相机的超级防抖功能确保了视频画面的稳定性,为运动员个人技术分析、教练训练计划制定以及比赛回顾提供了高质量的影像资料。

#### 1.4 眼球运动的类型

眼球运动包括平滑追踪(追随)、快速眼跳(扫视)和注视。平滑追踪可以帮助运动员更好地跟踪飞行中的网球,从而提高运动员对球的落点和速度的预判准确性。快速眼跳可以帮助运动员提升在不同视觉目标间切换的速度和准确性,有利于运动员更快地评估全场态势,做出更及时的战术调整。注视则可以帮助运动员在注视网球的过程中随时关注对手的击球姿势和站位,进而有效提升运动员在比赛过程选择对战术和预判对手运动的准确性。

#### 1.5 主要观察指标

主要观察指标包括:① 网球运动员眼球运动捕捉轨迹;② 真实运动轨迹;③ 深度学习模型损失值的变化。

#### 1.6 研究方法

在现实网球比赛过程中,由于运动员注意力的高度集中,其眼球的变化较为迅速和细微。为避免环境造成的网球运动员视频图像干扰,在采集网球运动员视频图像时,选择清晰度较高、帧率较快的设备。但该方法仅能降低比赛现场环境造成的误差,无法直接展示运动员的眼球轨迹。同时考虑到眼球在所采集的整个图像中所占的面积比例,直接使用视频图像阈值分割技术会导致所分割的图像中眼球的展示模糊不清。而 Tsallis 相对熵可以根据图像中的灰度值,通过图像中不同区域的灰度差,有效地分割出眼眶所在区域<sup>[11]</sup>。由此,更便于识别眼球运动的特征。

然而,眼球的运动特征较为复杂,无法通过数据直接表明不同特征。卷积神经网络可以通过学习运动员的眼球运动轨迹数据,不断迭代,以提高卷积神经网络在眼球运动特征识别过程中的准确性。引入注意力机制,使得眼球中的运动特征可以被更准确地识别,以保证眼球运动特征的识别效果。此外,考虑到卷积神经网络和注意力机制结合使用过程中会增加算法的复杂性,导致运行过程中出现过拟合现象,影响眼球运动轨迹捕捉的准确性。正则化可以通过损失函数优化卷积神经网络和注意力机制在处理眼球运动特征过程中的复杂度,有效降低算法出现过拟合现象的概率,不仅保证了算法对现有眼球运动图像的认识准确性,还可提高算法对未知图像的处理能力,保证算法的有效性。

##### 1.6.1 网球运动员视频图像多阈值分割

多阈值分割方法根据图像的不同区域设置不同的阈值,适应不同光线条件、背景干扰等复杂场景,从而更准确地分割出眼球区域<sup>[12-13]</sup>。Tsallis 相对熵作为多阈值分割方法中一种独特的度量方式,用于衡量两个概率分布之间的差异,且该度量具有非对称特性<sup>[14-15]</sup>。图像中的像素灰度值具有一定的分布规律,不同的灰度值在图像中的出现频率有所不同。因此,在图像分割的具体操作流程中,将图像中像素的灰度值视为一个概率分布。

通过计算不同区域之间的 Tsallis 相对熵,将图像中不同部分(如眼球区域和其他背景区域)的像素灰



图 2 网球运动员视频图像

度值看作不同的概率分布, 进而计算眼球区域像素灰度值概率分布与其他背景区域像素灰度值概率分布之间的 Tsallis 相对熵<sup>[16]</sup>。通过上述方式, 能够更加细致、准确地将眼球区域从复杂的图像中区分出来。Tsallis 相对熵实现网球运动员视频图像多阈值分割的详细运算过程如下:

令  $\mathbf{A}=(a_1, a_2, \dots, a_n)$ 、 $\mathbf{E}=(e_1, e_2, \dots, e_n)$  分别表示网球运动员视频图像内两个离散概率向量, 计算这两个离散概率向量的 Tsallis 相对熵测度  $\Phi(\mathbf{A}|\mathbf{E})$ , 其表达式如下:

$$\Phi(\mathbf{A}|\mathbf{E})=\frac{\left[\sum_{i=1}^n e_i\left(\frac{a_i}{e_i}\right)^q-1\right]}{q-1} \quad (1)$$

式中:  $q$  为网球运动员视频图像内离散概率向量总数。

在网球运动员视频图像的一维分布直方图内<sup>[17]</sup>, Tsallis 相对熵的灰度等级  $i$  的类概率表示为:

$$p_j^i=\frac{\exp\left[-\frac{(i-\mu)^2}{2\sigma_j^2}\right]\Phi(\mathbf{A}|\mathbf{E})}{\sqrt{2\pi\sigma_j}} \quad (2)$$

式中:  $\sigma_j$  为 Tsallis 相对熵的灰度等级  $i$  属于类别  $j$  的方差,  $\mu$  为 Tsallis 相对熵的灰度等级  $i$  的分布均值。

通过计算 Tsallis 相对熵的灰度等级  $i$  的类概率<sup>[18]</sup>, 建立 Tsallis 相对熵的灰度等级  $i$  的概率分布矩阵  $\mathbf{R}$ 。

令  $U$  表示网球运动员视频图像灰度等级直方图分布, 将  $R$  和  $U$  代入到公式(1)内, 得到关于  $R$  和  $U$  的 Tsallis 相对熵测度  $\Phi(U|R)$ , 在此基础上, 计算网球运动员视频图像多阈值分割的最佳阈值  $\xi$ , 其表达式如下:

$$\xi=\operatorname{argmin}[p_j^i\Phi(U|R)] \quad (3)$$

使用公式(3)计算得到的最佳阈值对网球运动员视频图像进行多阈值分割。最终得到网球运动员视频图像的前景图像, 以清晰展示运动员的轮廓和动作。由此可见, 通过运用 Tsallis 相对熵实现视频图像多阈值分割, 可以更快定位网球运动员的眼球区域, 减少不必要的计算和数据处理量, 有助于提高捕捉系统的整体效率, 使其能够更快地响应并实时记录眼球运动轨迹。

### 1.6.2 网球运动员眼球运动轨迹特征提取

多阈值分割虽然能够细致地划分图像中的不同区域, 但眼球运动轨迹数据包含大量信息, 如注视点位置、注视时长、眼跳幅度等, 且这些数据在不同情境下(如不同比赛场景、不同运动员状态)具有高度的复杂性和多样性<sup>[19-20]</sup>, 导致分割结果可能出现边缘模糊、区域重叠或遗漏等问题, 难以准确识别出网球运动员的轮廓和特征。

卷积神经网络(Convolutional Neural Networks, CNN)作为一种强大的深度学习模型(图 3), 在数据处理领域发挥着重要作用。CNN 的核心运算为卷积运算, 能够从原始数据中自动挖掘并提炼出有效的特征<sup>[21-22]</sup>。当运动员观察不同目标或者场景时, 眼球运动轨迹会呈现出不同模式, CNN 能够从复杂的轨迹数据中发现规律, 进而提取出与运动员视觉注意力焦点以及认知状态相关的特征。

网球运动员在比赛过程中, 眼球的运动轨迹会随网球的运动轨迹而变化。当网球平移时, 眼球也相应进行小幅度的平移; 当网球距离运动员较近时, 运动员的眼球会更加聚焦于网球, 使得眼球出现旋转和缩放等现象。这些细小的运动变化会增加卷积神经网络的学习时长, 影响眼球运动轨迹捕捉的实时性。而注意力机制作为一个重点选择器, 可以专注于众多眼球运动特征中的重要部分, 以提高运动特征的提取效率, 并计算不同眼球运动特征的权重, 有效筛选出对眼球运动轨迹捕捉过程中更为关键的特征, 从而提高运动轨迹捕捉的实时性。由此, 将网球运动员连续动作前景图像作为输入, 由卷积神经网络和注意力机制

组成深度学习神经网络模型捕捉网球运动员眼球运动轨迹。

当网球运动员的连续动作前景图像输入到深度学习神经网络模型时, CNN 通过局部感受野的方式对输入图像的局部区域进行感知, 并使用多个卷积核对输入图像进行卷积, 从而提取图像中的局部特征。

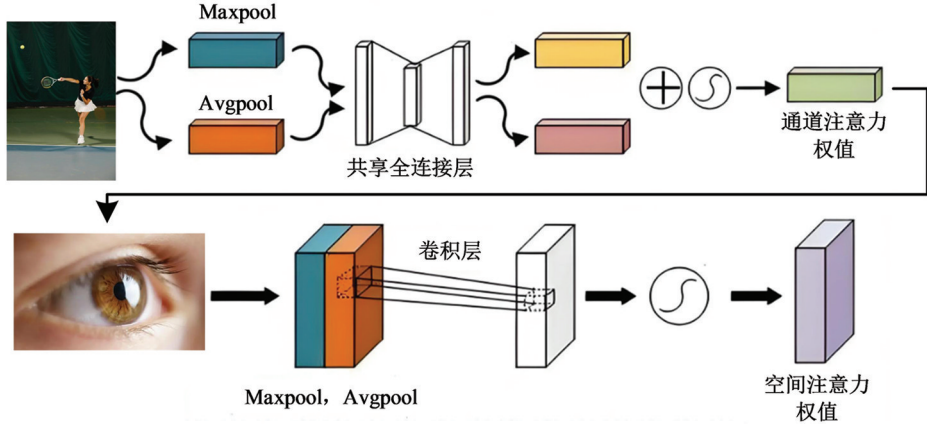


图 3 深度学习神经网络模型结构

输入为网球运动员连续动作分割后得到的前景图像序列, 图像尺寸为  $W \times H \times C$ , 其中  $W$  是宽度,  $H$  是高度,  $C$  是颜色通道数。使用多个卷积核  $K_i$  对输入图像进行卷积操作, 每个卷积核尺寸为  $k_w \times k_h \times C$ , 其中  $k_w$  和  $k_h$  分别是卷积核的宽度和高度。卷积操作公式为:

$$Y_i = I \times K_i + b_i \quad (4)$$

式中:  $I$  是输入图像;  $Y_i$  是第  $i$  个卷积核的输出特征图;  $b_i$  是偏置项。

对卷积层的输出特征图进行池化操作, 以减少特征图的尺寸并降低计算量。即在特征图上使用滑动窗口, 选择窗口内的最大值作为输出。将卷积层和池化层提取的特征图展平, 输入到全连接层中, 进行特征融合和分类。将网球运动员眼球位置轨迹通过编码转换成统一长度后<sup>[23]</sup>, 输出网球运动员眼球位置轨迹特征向量, 由  $(v_1, v_2, \dots, v_n)$  表示。

将卷积神经网络得到的结果输入到深度学习神经网络模型的注意力机制内<sup>[24]</sup>。注意力机制负责动态调整深度学习神经网络模型, 使深度学习神经网络模型在每个时间步均可依据当前运算需求关注输入网球运动员眼球位置轨迹特征向量的不同部分, 同时输出网球运动员眼球运动轨迹捕捉结果。注意力机制网络运行过程为:

将  $(v_1, v_2, \dots, v_n)$  输入到注意力机制网络内, 该网络在不同通道上对  $(v_1, v_2, \dots, v_n)$  进行挤压操作, 将输入特征图的高维信息整合到低维空间中, 从而提高计算效率<sup>[25]</sup>。由此, 生成含有注意力的网球运动员眼球运动轨迹特征  $\varphi$ , 其表达公式如下:

$$\varphi = f_{sq}(v_1, v_2, \dots, v_n) \quad (5)$$

式中:  $f_{sq}$  表示 Squeeze 压缩变换函数。

将公式(5)结果输入到注意力机制网络的 2 个全连接层并通过全连接层拟合不同通道网球运动员眼球运动轨迹特征, 得到网球运动员眼球运动轨迹特征的权重, 使得模型在面对不同场景和条件时, 能够更加准确地识别网球运动员的眼球运动轨迹。网球运动员眼球运动轨迹特征的权重  $H$  表达公式如下:

$$H = f_{ex}(\varphi, \bar{\omega}) \quad (6)$$

式中:  $f_{ex}$  表示轨迹函数;  $\bar{\omega}$  为网球运动员眼球运动轨迹特征在通道内的宽度。

结合公式(5)、(6)得到最终的网球运动员眼球运动轨迹特征  $Z$ , 其表达公式如下:

$$Z = f_{scale}(\varphi, H) \quad (7)$$

式中:  $f_{\text{scale}}$  为 Scale 操作函数。

Scale 操作函数可以将数据按比例缩放至特定范围。考虑到不同的眼球运动轨迹特征具有不同的量纲和取值范围, 通过 Scale 操作可以将不同的眼球运动轨迹特征转换到相同的尺度上, 从而消除量纲对数据分析和建模的影响, 以提高算法的性能和收敛速度。

将公式(7)结果输入到注意力机制网络的全连接层内<sup>[26-27]</sup>并通过 softmax 函数输出网球运动员眼球运动轨迹捕捉结果  $y$ , 其公式如下:

$$y = \text{softmax}(\omega'Z + b) \quad (8)$$

式中:  $\omega'$ 、 $b$  为全连接层权重和偏置。

综上所述, 该方法通过 CNN 的局部感受野和多个卷积核对输入图像进行特征提取, 利用池化操作减少特征图尺寸并降低计算量。将提取的特征图展平并输入到全连接层中进行特征融合和初步分类。在此基础上, 引入注意力机制, 通过动态调整模型对不同部分的关注程度, 使模型能够依据当前运算需求关注输入特征向量的关键部分, 从而提高眼球运动轨迹捕捉的准确性和鲁棒性。

### 1.6.3 眼球运动轨迹捕捉优化

提取网球运动员的眼球运动轨迹特征后, 能够实现网球运动员的眼球运动轨迹捕捉。但是, 深度学习网络模型在实际应用过程中, 随着训练样本数量的逐步增多以及训练次数的不断增加, 模型的复杂度呈现上升趋势, 容易引发过拟合。在捕捉网球运动员眼球运动轨迹时, 过拟合会降低捕捉结果的准确性<sup>[28-29]</sup>。正则化是深度学习中防范过拟合的重要手段, 通过向损失函数中添加正则化项, 可以限制模型参数的取值范围, 避免模型对训练数据过度拟合<sup>[30]</sup>。对于由卷积神经网络和注意力机制组成的深度学习网络模型, 在正则化处理时, 会约束模型参数的范数, 进而降低深度学习网络模型处理网球运动员眼球运动轨迹特征的复杂度, 提高特征提取的实时性。

令  $p(\theta)$  表示深度学习神经网络模型的损失函数,  $\theta$  为模型内所有参数,  $J(\omega')$  表示深度学习神经网络模型的复杂度函数,  $\omega'$  为模型权重, 正则化是对  $p(\theta)$  和  $J(\omega')$  进行综合优化, 同时使用 L1 正则化和 L2 正则化约束, 优化深度学习神经网络模型的复杂度和损失函数。其中: L1 正则化通过引入权重的绝对值之和作为惩罚项, 促使模型产生稀疏的权重矩阵, 有助于特征选择并降低模型对训练数据的依赖; L2 正则化则通过引入权重的平方和作为惩罚项, 使模型产生平滑的权重, 提高模型的稳定性和鲁棒性。这种组合正则化方法不仅有助于模型在训练过程中更加关注对输出有显著影响的特征, 还能有效降低模型的复杂度, 防止过拟合, 从而提高模型的泛化能力。正则化参数的选择对模型性能有着关键影响: 若取值过大, 模型的复杂度会被过度抑制, 可能导致模型欠拟合, 无法充分学习数据中的有效特征, 使得对网球运动员眼球运动轨迹捕捉的准确性下降; 若取值过小, 难以对模型参数进行有效约束, 过拟合问题依旧存在。在实际应用中, 需通过实验调试, 采用交叉验证等方法, 在训练集和验证集上反复测试不同的正则化参数组合, 在捕捉网球运动员眼球运动轨迹任务中探索最优泛化能力的参数设置, 确保模型在面对新数据时能够稳定、准确地输出眼球运动轨迹捕捉结果。深度学习神经网络模型复杂度和损失函数优化函数的表达式如下:

$$O = |P(\theta)|_1 + \lambda |J(\omega')|_2 \quad (9)$$

式中:  $|P(\theta)|_1$ 、 $|J(\omega')|_2$  分别为 L1 正则化和 L2 正则化算子;  $\lambda$  为深度学习神经网络模型复杂度导致的损失占总损失的比重。利用公式(9)实现深度学习神经网络模型复杂度和损失函数的优化, 使其输出的网球运动员眼球运动轨迹捕捉结果更加精准。基于上述处理, 可以实现深度学习神经网络模型的优化, 有效避免模型在运行过程中由于特征的复杂性导致的过拟合, 降低模型对现有眼球运动轨迹数据的依赖程度, 提高

了模型捕捉运动员眼球运动轨迹的实时性和精准度。

## 2 结果与分析

### 2.1 实验设置

在现有案例的基础上,采用所提方法对网球运动员眼球运动轨迹捕捉过程进行优化。实验过程中使用的计算设备如表 2 所示。

表 2 计算设备列表

设备名称	硬件	型号
训练服务器	CPU	Intel Xeon E5-2680 v4 @ 2.40 GHz(14 cores)
	GPU	4 x NVIDIA Tesla V100(32 GB each)
	RAM	256 GB DDR4
	存储	4 TB NVMe SSD
实时处理工作站	CPU	Intel Core i9-11900K
	GPU	NVIDIA GeForce RTX 3090
	RAM	64 GB DDR4
	存储	1 TB NVMe SSD

实验步骤如下:

步骤 1: 数据采集。利用高速摄像机(如 GoPro HERO12 Black, 4K 分辨率, 120 fps)记录 20 名网球运动员在多种不同比赛场景(包括硬地、草地等不同场地类型,以及单打、双打等不同比赛形式)和训练环节(如日常基础训练、模拟对抗训练等)下的运动情景。同时使用 Tobii Pro Glasses 3 眼动仪采集眼球运动数据,将其作为精确的对比基准数据。选取的运动员涵盖了不同年龄阶段、不同性别、不同比赛经验水平以及不同惯用手的个体差异,最大程度地保证数据能够反映网球运动员群体在各类场景下的真实表现,从而增强后续研究结果在整个网球运动员群体中的普适性。

步骤 2: 数据预处理。

- 1) 视频帧提取: 使用 OpenCV 库从视频中提取图像帧。
- 2) 数据标注: 利用专业软件对部分数据进行人工标注,标记眼球位置和运动轨迹。
- 3) 数据增强: 通过旋转、缩放、添加噪声等方法扩充数据集。

步骤 3: 数据集划分。将数据集按 6: 2: 2 的比例划分为训练集、验证集和测试集。

步骤 4: 多阈值分割。

- 1) 实现 Tsallis 相对熵计算函数,根据公式(1)和(2)计算图像的灰度等级概率分布。
- 2) 利用公式(3)计算最佳阈值。
- 3) 使用确定的最佳阈值对网球运动员视频图像进行多阈值分割,得到前景图像。

步骤 5: 深度学习模型构建阶段。

- 1) 设计 CNN 架构,包括卷积层、池化层和全连接层。实现公式(1)描述的卷积操作。
- 2) 实现注意力机制,包括 Squeeze 操作(公式(3))和 Excitation 操作(公式(4))。集成 Scale 操作(公式(5))和最终的 softmax 输出层(公式(6))。

步骤 6: 模型优化。实现 L1 和 L2 正则化(公式(7))。

设置实验过程中算法的参数如表 3 所示。

表 3 实验参数设置

参数	参数值
初始阈值数量/个	3~5
迭代次数/次	1 000
正则化参数	0.01
CNN 基础网络	VGG16
输入图像尺寸/像素	5 568×4 872
批次	32
学习率	0.001
正则化项	L1: 0.000 1, L2: 0.000 5
训练轮数/轮	50

在上述设置的基础上,对网球运动员视频图像进行多阈值分割,获取网球运动员视频图像的前景是捕捉网球运动员眼球运动轨迹的基础。以某帧网球运动员视频图像作为实验对象,使用本文方法对其进行多阈值分割,结果如图 4 所示。

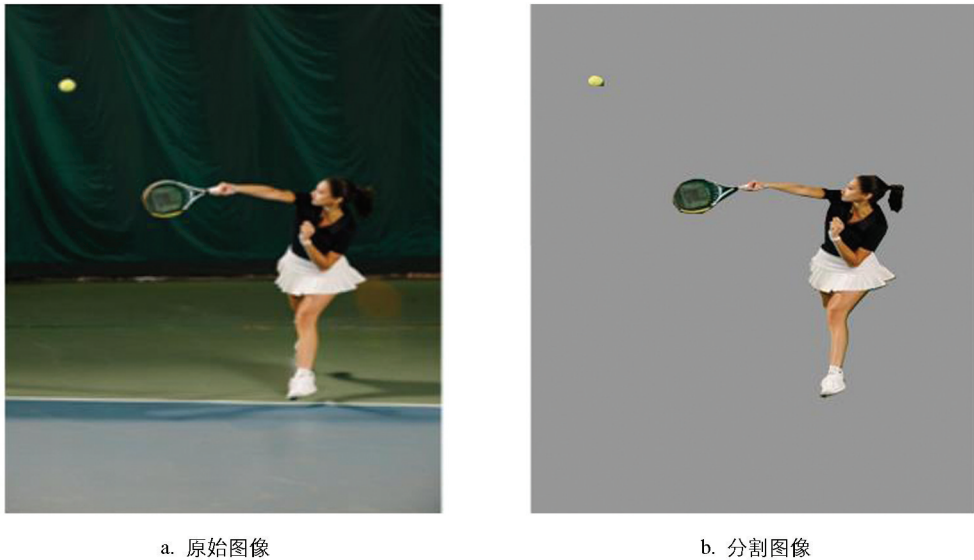


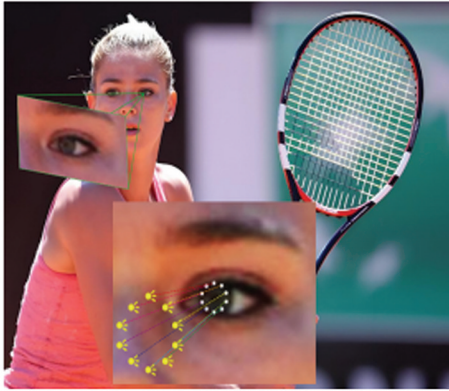
图 4 网球运动员视频图像多阈值分割测试结果

如图 4 所示,所提方法在网球运动员视频图像的处理过程中,能够精准地区分运动员与复杂多变的背景,有效提取运动员打网球时的清晰前景图像,并保留了关键的运动细节和轮廓信息。这是因为所提方法采用 Tsallis 相对熵方法进行视频图像多阈值分割,相比单一阈值分割能够更精细地处理图像中的不同灰度级,使得分割出的前景图像更加清晰,同时保留了关键的运动细节和轮廓信息。该方法能够自适应地确定多个阈值,从而更准确地分割出网球运动员与复杂多变的背景,充分验证了本文方法的高效性与准确性,为后续深入分析提供了高质量的图像素材。精确的前景分割能够减少背景噪声干扰,提升眼球运动轨迹捕捉的准确性。

## 2.2 结果分析

以 4 帧国际赛事网球运动员视频图像作为分析对象,使用本文方法捕捉图像内运动员的眼球运动轨迹,捕捉结果如图 5 所示。本文方法在网球运动员视频图像中成功捕捉到眼球,并准确地标记出网球运动员眼球中的网球,直观展示了运动员的视线变化。这是因为本研究能够利用 CNN 多层卷积和池化操作逐步实现从低级特征(如边缘、纹理)到高级特征(如眼球的形状和位置)的提取。这种层级特征提取方式非常适合捕捉眼球在图像中的复杂表现。同时,池化操作赋予了 CNN 平移不变性,使其能够识别出位于图像任

意位置的眼球,这对于追踪运动中的眼球至关重要。



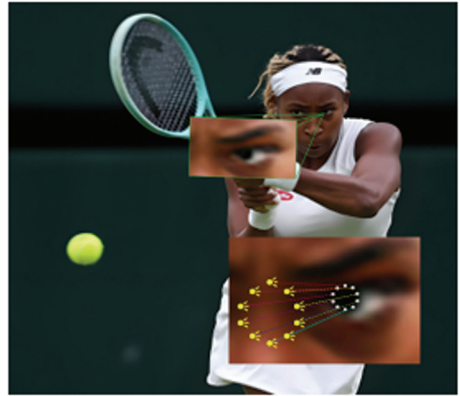
a. 注视



b. 平滑追踪



c. 快速跟跳



d. 平滑追踪

图 5 网球运动员眼球轨迹捕捉结果示例

连续性网球运动员眼球运动轨迹捕捉结果如图 6 所示。从图 6 中可以看出,本文方法捕捉到的眼球运动轨迹与真实的网球运动员眼球运动轨迹高度吻合。这是因为本文方法通过 Tsallis 相对熵方法进行多阈值分割,有效地将运动员(前景)与背景分离。通过该步骤,能够提高后续分析的针对性和准确性,使得模型专注于包含眼球运动信息的关键区域。结合 CNN 和注意力机制的模型能够自适应地处理不同的场景、光照条件和眼球运动速度,提高模型在各种比赛或训练情况下的表现。因此,本研究方法的精确度较高,在实际应用中能够精准捕捉眼球运动的轨迹。

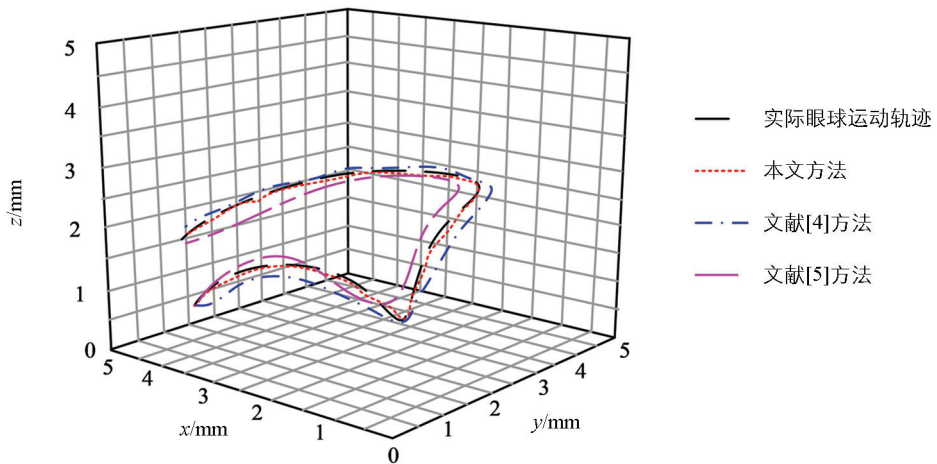


图 6 连续性网球运动员眼球运动轨迹捕捉结果

为验证本文方法对深度学习神经网络模型的优化能力,以深度学习神经网络模型训练过程中的损失值作为衡量指标,测试不同训练代数情况下深度学习神经网络模型损失值变化情况,测试结果如图 7 所示。

分析图 7 可知,本文方法使用 Tsallis 相对熵方法进行多阈值分割,不仅有效分离了前景(运动员)和背景,还保留了更多细节信息。这种高质量的预处理为深度学习模型提供了更清晰和更相关的输入数据,大大减少了模型需要学习的噪声和无关信息,从而加速训练过程并提高模型收敛速度。通过引入正则项,解决了深度学习模型可能出现的过拟合问题。因此,本文方法的损失值会随训练代数的增加而降低并逐渐趋于 0。在 3 种方法中,本文方法将深度学习神经网络模型损失值降至最低所需的训练代数最少。

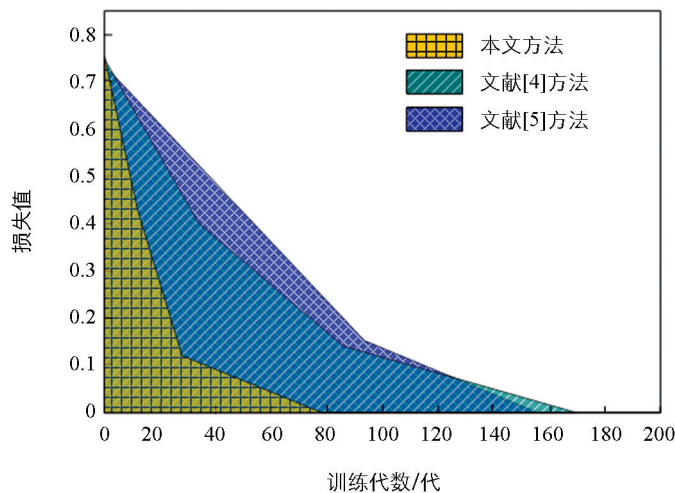


图 7 深度学习神经网络模型优化结果

为验证本文方法在捕捉网球运动员眼球运动轨迹方面的效果,以误差率为指标,测试结果如表 4 所示。

表 4 3 种方法的运动轨迹捕捉误差率对比

方法	实验视频 1 误差率/%	实验视频 2 误差率/%	实验视频 3 误差率/%	实验视频 4 误差率/%	实验视频 5 误差率/%	平均 误差率/%	误差率 标准差
本文方法	3.0	3.3	3.1	3.2	3.4	3.2	0.15
文献[5]方法	8.2	8.8	8.6	8.4	8.5	8.5	0.22
文献[4]方法	6.5	7.0	6.9	6.7	6.9	6.8	0.21

由表 4 可知,在 5 个不同的实验视频测试中,本文方法在捕捉网球运动员眼球运动轨迹方面的误差率较低,平均误差率仅为 3.2%,且误差率标准差为 0.15,表明其在不同视频中的误差波动较小,稳定性高。这是因为本文方法采用 Tsallis 相对熵进行多阈值分割,能够减少背景干扰对眼球运动轨迹捕捉的影响。基于卷积神经网络和注意力机制构建的深度学习模型能够自动挖掘和提炼眼球运动轨迹的有效特征,并根据不同场景动态调整对特征的关注程度,提高了特征提取的准确性和鲁棒性。因此,本文方法在网球运动员眼球运动轨迹捕捉方面展现出较好的性能,有效降低了误差率。

### 3 讨论与启示

网球运动员在比赛过程中产生的眼球运动轨迹是运动员在比赛过程中获得胜利的关键。因为细微的眼球运动可以帮助运动员精准捕捉到运动的网球,进而让运动员选择更加适合的比赛战术,以提高比赛获胜的概率。然而,现有研究中,对网球运动员眼球运动轨迹的捕捉研究较少,无法为运动员提供更有效的数据来支持运动员修改和提升训练计划,为此,提出视频图像阈值分割下的网球运动员眼球运动轨迹捕捉优化方法。相较于传统的基于简单阈值分割或单一模型的眼球运动轨迹捕捉方法,本文将 Tsallis 相对熵多阈值分割技术与卷积神经网络、注意力机制以及正则化优化相结合。Tsallis 相对熵<sup>[31]</sup>能够更为精准地处理图像中复杂的灰度分布,为后续深度学习模型提供高质量的输入。卷积神经网络<sup>[32]</sup>与注意力机制适用于处理具有高度复杂性和动态性的数据特征,能够使模型在复杂多变的网球比赛场景下自适应地提取眼球运动的关键特征。该方法为运动视觉领域的研究提供了新的技术思路,有助于深入理解网球运动员视觉认知与决策机制。

实验结果表明,所提方法可以有效地分割出网球运动员视频图像中的目标,并捕捉到清晰的网球运动

员眼球运动轨迹。这是因为所提方法利用视频图像分割技术中的 Tsallis 相对熵, 基于图像不同区域的灰度差, 较为精确地分割出所采集图像中的运动员眼眶部分, 并将其输入至卷积神经网络中, 提取眼球运动轨迹。同时, 利用注意力机制提高眼球运动特征的清晰程度, 以保证眼球运动特征的提取效果。在此基础上, 采用正则化技术降低算法的复杂度, 保证运动轨迹捕捉的实时性。

此次研究存在一些局限性: 在数据采集方面, 虽然选择了特定型号运动相机并设置了相关参数, 但仍未涵盖所有复杂的网球运动场景和运动员个体差异情况。例如: 强光直射或极低光照对图像采集质量会产生不良影响; 不同水平或风格运动员的眼球运动特征多样性未被充分挖掘。实验对象仅选取了闽南师范大学体育学院网球运动员, 样本范围相对较窄, 可能存在一定的样本偏差, 无法完全代表所有网球运动员群体的特征, 对研究结果的普遍性有一定影响。此外, 在模型构建方面, 尽管当前的深度学习模型在性能上表现出色, 但面对极端复杂场景下的网球运动员眼球运动, 模型的泛化能力仍有待提升。同时, Tsallis 相对熵计算过程较复杂, 在一定程度上影响了处理速度, 这在实时性要求极高的场景中可能成为限制因素。

未来研究可从以下方面展开。在数据采集方面, 进一步探索更先进的图像采集设备和技术, 能够适应更广泛的环境条件, 获取更高质量、更全面的运动员视频图像数据, 同时结合多种传感器获取更多维度的信息, 以更精准地同步记录眼球运动与身体动作等相关数据, 为更深入的研究提供丰富的数据支持。后续研究可引入迁移学习技术, 利用大规模跨领域视觉数据对模型进行预训练, 使其能够更好地适应不同场景。对于 Tsallis 相对熵计算复杂度高的问题, 可以采用并行计算技术提高整体处理效率。还可拓展应用领域, 将该方法用于其他球类运动项目中, 研究不同运动项目运动员眼球运动特征差异及其与运动表现的关系, 建立更通用的运动员视觉行为模型, 为体育训练和竞技提供更广泛的理论支持和技术指导。同时, 也可探索将眼球运动轨迹捕捉技术与虚拟现实、增强现实技术相结合, 为运动员模拟训练、技能提升提供更沉浸式、个性化的训练环境和反馈机制。

## 参考文献:

- [1] 任条娟, 陈鹏, 陈友荣, 等. 基于深度学习的多目标运动轨迹预测算法 [J]. 计算机应用研究, 2022, 39(1): 296-302.
- [2] 朱琳, 董鹏程, 沈培鑫, 等. 基于 SGGC-Net 动作捕捉系统解析行走步态的信度与效度 [J]. 医用生物力学, 2024, 39(2): 305-311.
- [3] 骆蕾, 张杨, 陈洁, 等. 发作性睡病患者的快速眼球运动密度特征 [J]. 中国心理卫生杂志, 2022, 36(8): 658-662.
- [4] 支慧晶, 刘阳. 竞走裁判员在多人情景下判罚决策的视觉搜索特征 [J]. 心理与行为研究, 2024, 22(1): 130-136.
- [5] 胡启慧, 蔡英凤, 王海, 等. 基于层次图注意的异构多目标轨迹预测方法 [J]. 汽车工程, 2023, 45(8): 1448-1456.
- [6] 李新, 陈伟炯, 陈文涛, 等. 基于眼动追踪技术的引航员情景意识识别研究 [J]. 安全与环境学报, 2024, 24(4): 1485-1494.
- [7] 陈静, 黎雅丽, 陆泉. 基于眼动追踪的用户感知预测模型研究 [J]. 情报理论与实践, 2022, 45(4): 154-161, 169.
- [8] 翁冬冬, 王怡晗, 郭署山, 等. 融合时空约束的光学动作捕捉标记点实时补全方法 [J]. 计算机辅助设计与图形学学报, 2023, 35(8): 1197-1205.
- [9] 杨蕴, 李玉, 赵泉华. 基于局部空间信息的可变类模糊阈值光学遥感图像分割 [J]. 自动化学报, 2022, 48(2): 582-593.
- [10] 吴佳芸, 武灵芝, 胡晓飞. 基于通用学习均衡优化器的多阈值图像分割 [J]. 传感技术学报, 2024, 37(3): 463-468.
- [11] KANG L Q, YANG P C, LÜ Q J, et al. Wrapping Phase Repair Method Based on Tsallis Relative Entropy Evaluation [J]. Optics and Precision Engineering, 2024, 32(8): 1130-1139.
- [12] 董维振, 陈燕. 基于改进被囊群算法的图像多阈值分割方法 [J]. 计算机工程与设计, 2023, 44(7): 2093-2102.
- [13] 刘庆鑫, 李霓, 贾鹤鸣, 等. 改进鲫鱼优化算法和熵测度的图像多阈值分割 [J]. 智能系统学报, 2024, 19(2): 381-391.

- [14] 康乐谦, 杨鹏程, 吕秋娟, 等. Tsallis 相对熵引导的包裹相位修补 [J]. 光学精密工程, 2024, 32(8): 1130-1139.
- [15] 方叶祥, 杜和祥, 甘平. 基于 Tsallis 相对熵的 DEMATEL 阈值确定研究 [J]. 工程数学学报, 2023, 40(2): 207-218.
- [16] 汤弘毅, 徐武, 杨昊东, 等. 基于 IWOA 算法的 Tsallis 相对熵图像多阈值分割 [J]. 江苏科技大学学报(自然科学版), 2022, 36(1): 90-97.
- [17] 丁红发, 傅培旺, 彭长根, 等. 混洗差分隐私保护的度分布直方图发布算法 [J]. 西安电子科技大学学报(自然科学版), 2023, 50(6): 219-236.
- [18] 张昱彤, 翟旭平, 聂宏. 基于低分辨率红外传感器的深度学习动作识别方法 [J]. 红外技术, 2022, 44(3): 286-293.
- [19] 张欣宇, 邱国鹏. 融合改进 LM 算法及动态时间规整算法的人体动作捕捉研究 [J]. 西南大学学报(自然科学版), 2024, 46(5): 175-185.
- [20] 王雨婷, 梁旭鹏, 许国良, 等. 基于混合运动激励和时序增强的篮球运动员动作识别算法 [J]. 重庆邮电大学学报(自然科学版), 2024, 36(2): 307-318.
- [21] WANG X Y, LIU J. Spatially Regularized Leaky ReLU in Dual Space for CNN Based Image Segmentation [J]. Inverse Problems and Imaging, 2024, 18(6): 1320-1342.
- [22] 赵凤, 耿苗苗, 刘汉强, 等. 卷积神经网络与视觉 Transformer 联合驱动的跨层多尺度融合网络高光谱图像分类方法 [J]. 电子与信息学报, 2024, 46(5): 2237-2248.
- [23] 许才项, 贺杰, 庞家豪. 基于卷积神经网络的分形图像编码研究 [J]. 计算机仿真, 2024, 41(8): 210-214.
- [24] 刘悦, 张雷, 辛山, 等. 融入时空注意力机制的深度学习网络视频动作分类 [J]. 中国科技论文, 2022, 17(3): 281-287.
- [25] 李昊, 曹书瑜, 陈亚青, 等. 基于注意力机制的用户轨迹识别模型 [J]. 计算机科学, 2022, 49(3): 308-312.
- [26] 陈广秋, 尹文卿, 温奇璋, 等. 基于双重注意力机制生成对抗网络的偏振图像融合 [J]. 电子测量与仪器学报, 2024, 38(4): 140-150.
- [27] 夏晓华, 钟预全, 胡鹏, 等. 综合多尺度信息和注意力机制的水下图像增强 [J]. 光学精密工程, 2024, 32(10): 1582-1594.
- [28] 王冲, 魏子令, 陈曙晖. 基于自注意力机制的无边界应用动作识别方法 [J]. 计算机研究与发展, 2022, 59(5): 1092-1104.
- [29] 杜芳芳, 王福忠, 高继梅. 空洞卷积网络下微表情实时识别方法仿真 [J]. 计算机仿真, 2023, 40(7): 172-175, 461.
- [30] 陈善学, 许少华. 基于图拉普拉斯正则化的柯西非负矩阵分解高光谱解混 [J]. 激光与光电子学进展, 2024, 61(14): 278-288.
- [31] BEHERA S, CONTRERAS-REYES J E, KAYAL S. Mutual Information Matrix and Global Measure based on Tsallis entropy [J]. Nonlinear Dynamics, 2025, 113(6): 5239-5249.
- [32] 王悦天, 傅司超, 彭勤牧, 等. 半监督场景下多视角信息交互的图卷积神经网络 [J]. 软件学报, 2024, 35(11): 5098-5115.

责任编辑 张枸  
柳剑