

Segment Anything 在多人运动中的应用

综述：基于一个可以精准切割图片中各部分的网络 Segment Anything (SAM)，本文希望探讨其在多人运动领域的应用，以解决以往的数据库受到运动员遮挡等问题而出现的识别错误。本文打算基于 SAM 建立一个综合分析系统，由用户自定义分类运动员的标准后输入视频录播，在 SAM 切割出运动员并添加标签后，该系统可以根据对应模块实现运动员识别及标记、动作识别及评估、重识别、生成对特定运动员的报告与高光时刻集锦、比赛犯规情况判定等五种功能。该系统由于使用了 SAM，精准度可以大幅提升，同时由于 SAM 是已训练好的泛用性强的网络，训练该网络花费的成本也会减少。

1. 介绍

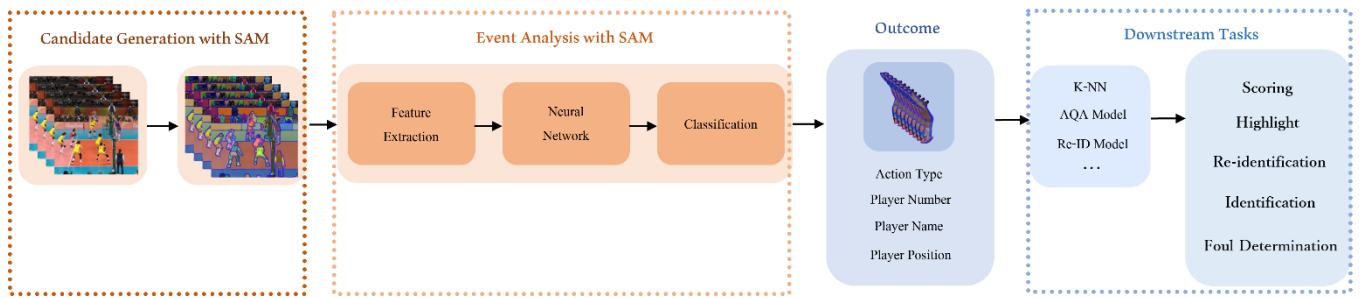
在计算机视觉领域中，图像分割是一项重要任务，旨在将图像中的像素分配给不同的语义类别。虽然已有多种方法能够在图像中识别和分割出不同的对象，但对于多人运动的场景，目前仍存在一定的挑战。常见的多人运动包括足球、排球和篮球等，场上运动员会发生遮挡、交叉等情况，这些因素会影响对象的识别和跟踪。

当前已经有一些多人运动数据集，比如 MultiSports 可以对不同运动场景中的运动员进行识别 [1] (Li 等, 2021)，MPS-Net 也可以对运动员的 3D 模型进行构建 [2] (Wei 等, 2022)。但是 Multisports 只能根据 bounding box 来识别对象，无法实现精准的图像分割。此外，由于多人运动中存在运动员交叉问题，使得 MPS-Net 等网络识别的精准度也有所降低。

为解决这些问题，研究人员最近提出了一种 Segment Anything (SAM) 的方法，该方法能够实现对多人运动场景中的对象进行精准的分割。SAM 是一个高效的模型，通过在数据收集循环中使用该模型，研究人员建立了迄今为止最大的多人运动数据集，包括超过 11M 张图片和 10B 个 mask [3] (Kirillov 等, 2023)。

SAM 是一种 promptable 的模型，可以进行零样本学习，能够将已学习到的知识迁移到新的图像分布和任务中，且其零样本学习的性能表现卓越，通常能够与或甚至超越以前完全监督的结果。在多人运动场景中，SAM 可以实现对场上运动员的精准切割，并能够对其进行动作识别、动作评估、球员重识别、球衣号码识别等任务，具有较为广阔的应用前景。

2. 方法



基于 SAM 系统，我们可以自动化地将多人运动视频按照运动员分割成多个姿势序列，为多人运动数据的分析和理解提供了新的手段。

在该系统中，我们设计了一些使用者输入指令，用于根据用户需要自定义运动切割方式。首先，用户可以在系统中选择使用预设的运动切割方案，比如“Soccer”、“Basketball”和“Volleyball”等，这些方案方便调用已有的分析系统。同时，使用者可以给定切割后的分类方法，如按照运动员的球衣号码进行分类，按照运动场上的位置进行分类等。随后，输入比赛视频，SAM 可以准确地切割各运动员的边框，识别出他们的位置。

在运动员切割完成后，系统基于 SAM 切割运动员的结果提供了五种可选的功能。

(1) 运动员识别及标记。根据使用者提供的分类依据，系统可以根据要求，利用 SAM 切割出对应球员的球衣号码、位置等信息，然后采用 k-NN 等算法，识别出图像中球员的上述信息[4] (Wang & Sharma, 2022) 并提供给切割运动员相应的标签，用于进一步的分析。

(2) 动作识别及评估。我们打算基于常见的篮球、足球、排球中的动作对我们的神经网络加以训练，根据多人运动的类型以及提供的动作类型，该系统可以识别运动员的动作，如在足球比赛中识别射门等。该识别结果会用于给运动员添加标签以分类，便于后续功能的处理。

基于识别出的动作和打分的模型，我们可以实现对运动员动作的打分与评估，并且会统计运动员执行指定动作的次数以及动作的成功率。我们打算通过识别切割出来的运动员的关节点位置评估运动员的动作，借助已有的 Action Quality Assessment 模型[5] (Pan 等, 2019)，通过对其在多人运动技术动作方面的训练，使得我们的系统具备给制定运动员动作打分的功能。

(3) Re-identification. 在球类比赛录播中通常会涉及到摄像机视角转换等问题，需要对运动员进行重识别。我们打算根据视角转换前后 SAM 的切割结果，若转换视角前后球衣号码识别结果相同，则认为该运动员为同一运动员。

(4) 生成对特定运动员的报告与高光时刻集锦。针对一场特定比赛，我们打算根据运动员的动作评估等数据分析，生成其关于此场比赛的表现报告。同时，系统使用者可以设定分数界限，分数超过该界限的该运动员的动作会被提取出来，再根据这些动作对应视频片段，生成该运动员的高光视频。

(5) 比赛犯规情况判定。可以通过识别运动员的位置、动作等信息，判断该运动员是否犯规。该应用已被用在世界杯足球赛中[6] (Held 等, 2023)，应用 SAM 可以将判断是否犯规的应用范围进一步扩展到篮球、排球等比赛方面。

3. 应用前景

首先，使用 SAM 进行运动员识别可以提高识别的精准度。传统的多人运动数据集往往只能根据 bounding box 来识别，而使用 SAM 可以实现对运动员的精准切割，解决遮挡等问题，从而提高识别的准确率。

其次，SAM 是一种已有模型，训练简便。它使用的是预训练的图像分割模型，不需要大量的样本数据进行训练，因此可以节省训练时间和成本。

此外，使用 SAM 进行运动员识别还可以实现动作识别、动作评估、比赛双方胜率预测、高光时刻生成等功能，这些功能有助于教练员更好地指导球员，提高球员的训练效果。

SAM 可以应用于足球、排球和篮球等多个运动项目。例如，在足球比赛中，SAM 可以实现对球员和裁判员的识别，以及犯规判断等功能；在排球比赛中，SAM 可以实现对运动员的识别和姿态分析，从而为球队制定更加科学的战术；在篮球比赛中，SAM 可以实现对球员的识别和球衣号码识别，从而为球员评分提供依据，同时也可以用于实时统计球员的表现情况等。

使用 SAM 进行多人运动中的运动员识别，可以提高识别精准度，训练成本较低，同时还能实现多种功能，未来在多人运动领域中的应用前景广阔。

References:

- 【1】Li, Y., Chen, L., He, R., Wang, Z., Wu, G., & Wang, L. (2021).

*MultiSports: A Multi-Person Video Dataset of Spatio-Temporally
Localized Sports Actions.* 13536 – 13545.

https://openaccess.thecvf.com/content/ICCV2021/html/Li_MultiSports_A_Multi-Person_Video_Dataset_of_Spatio-Temporally_Localized_Sports_Actions_ICCV_2021_paper.html

- 【2】 Wei, W.-L., Lin, J.-C., Liu, T.-L., & Liao, H.-Y. M. (2022). *Capturing Humans in Motion: Temporal-Attentive 3D Human Pose and Shape Estimation from Monocular Video* (arXiv:2203.08534). arXiv.
<https://doi.org/10.48550/arXiv.2203.08534>

- 【3】 Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2023). *Segment Anything* (arXiv:2304.02643). arXiv.
<https://doi.org/10.48550/arXiv.2304.02643>

- 【4】 Wang, L., & Sharma, A. (2022). Analysis of sports video using image recognition of sportsmen. *International Journal of System Assurance Engineering and Management*, 13(1), 557 – 563.
<https://doi.org/10.1007/s13198-021-01539-4>

- 【5】 Pan, J.-H., Gao, J., & Zheng, W.-S. (2019). Action Assessment by Joint Relation Graphs. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 6330 – 6339. <https://doi.org/10.1109/ICCV.2019.00643>

- 【6】 Held, J., Cioppa, A., Giancola, S., Hamdi, A., Ghanem, B., & Van

Droogenbroeck, M. (2023). *VARS: Video Assistant Referee System for Automated Soccer Decision Making from Multiple Views*

(arXiv:2304.04617). arXiv. <https://doi.org/10.48550/arXiv.2304.04617>