International Journal of Computer Vision manuscript No.
(will be inserted by the editor)

A Decade of Action Quality Assessment: Largest Systematic Survey of Trends, Challenges, and Future Directions

Hao Yin[1, 2*] · Paritosh Parmar[3*] · Daoliang Xu[1, 2] · Yang Zhang[2] · Tianyou Zheng[2] ✉ · Weiwei Fu[1, 2] ✉

Received: date / Accepted: date

Abstract Action Quality Assessment (AQA)—the ability to quantify the quality of human motion, actions, or skill levels and provide feedback—has far-reaching implications in areas such as low-cost physiotherapy, sports training, and workforce development. As such, it has become a critical field in computer vision and video understanding over the past decade. Significant progress has been made in AQA methodologies, datasets, and applications, yet a pressing need remains for a comprehensive synthesis of this rapidly evolving field. In this paper, we present a thorough survey of the AQA landscape, systematically reviewing over 200 research papers using the preferred reporting items for systematic reviews and meta-analyses (PRISMA) framework. We begin by covering foundational concepts and definitions, then move to general frameworks and performance metrics, and finally discuss the latest advances in methodologies and datasets. This survey provides a detailed analysis of research trends, performance comparisons, challenges, and future directions. Through this work, we aim to offer a valuable resource for both newcomers and experienced researchers, promoting further exploration and progress in AQA.

*These authors contributed equally to this work.

✉ Weiwei Fu
E-mail: fuww@sibet.ac.cn
✉ Tianyou Zheng
E-mail: zhengty@sibet.ac.cn
[1] School of Biomedical Engineering (Suzhou), Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui, China
[2] Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Science, Suzhou, Jiangsu, China
[3] Institute of High-Performance Computing, Agency for Science, Technology and Research, Singapore

Keywords action quality assessment, skills assessment, action understanding, video understanding, computer vision, deep learning, survey

1 Introduction

Skills and Action Quality Assessment (**AQA**) is an *emerging* and *critical* field in video understanding, moving beyond action recognition and action prediction [1] to **evaluate how well** actions are performed and score the skill level of performers (see Fig. 1). These techniques are essential in a range of **domains**, including *sports* [2, 3], *healthcare* [4–6], *fitness* [7, 8], *industrial training* [9], & *AI video content generation* [10] where accurate assessment of human actions/performance is crucial.

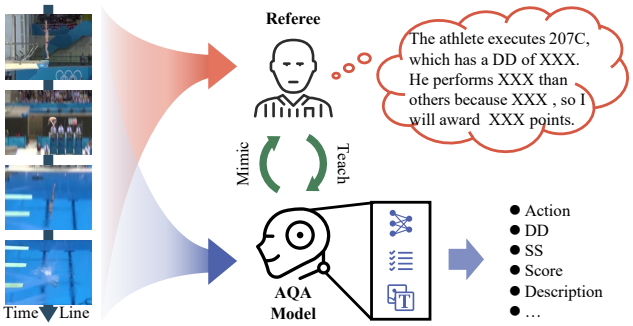

Fig. 1 AQA model plays the role of referee to evaluate how well actions are performed and score the skill level of performers.

For example, in *sports* [2, 3], it could be used to assess how well an Olympics diver performed, reporting what they did well/correctly, what they did wrong and what was the severity of these errors; take all (ideally)