

ABSTRACT

The present disclosure relates to a video processing method and an apparatus, an electronic device, and a storage medium. The method includes: performing a feature extraction on a plurality of target video frames of a video to be processed through a feature extraction network, to obtain feature maps of the plurality of target video frames; performing an action recognition process on the feature maps of the plurality of target video frames through an M-level action recognition network, to obtain action recognition features of the plurality of target video frames; and determining a classification result of the video to be processed according to the action recognition features of the plurality of target video frames. According to the video processing method of the embodiments of the present disclosure, the action recognition features of the target video frames may be obtained through a multi-level action recognition network, and the classification result of the video to be processed may be further obtained, without action recognition by a process such as optical flow or 3D convolution, reducing the amount of computation, improving the processing efficiency, allowing for online real-time classification on the video to be processed, and enhancing practicability of the video processing method.