Abstract

With the development of China's agricultural industry in the direction of automation, scale, and information, the traditional mode of small-scale labor and manual supervision has been unable to meet the growing demand of consumers for the quality of agricultural products. Applying machine vision based detection to the large-scale production of farms, we can get the information of each farmer's manual work without disturbing them. With the rapid development of artificial intelligence, researchers are not satisfied with the task of only accurate and quick target detection, but also the ability of autonomous learning. People like that computers can be more intelligent, can learn and understand the contextual relationships between multiple images, and thus understand the continuous behavior in visual system. In this paper, video analysis and understanding problems of FADers in farm are analyzed. Through deep learning and other technical method, singe working farmers are detected, based on which the spatiotemporal behavior features of farmers' labor are extracted. Finally, the automatic recognition of farmer working pattern is realized. The main research is as follows:

(1) Construction of farmer working dataset. Firstly, four kinds of farmer working pattern are defined, namely, pesticide, plowing land, weeding, and transplanting seedlings, denoted as PES, PLO, WEE, and TRA respectively. Massive farmer's labor videos are collected. Key frames of the video segments are extracted. By target detection and instance segmentation, video frame samples of single farmer's labor are generated. And FAD_Y dataset and FAD_YM dataset are built.

(2) Construction and improvement of C3D behavior recognition model. C3D network structure is built. FAD_Y and FAD_YM dataset being the input of C3D, a spatiotemporal feature extractor of farmers' labor behavior is generated. Finally, based on feature maps, classifiers, like softmax and SVM, are used to classify the labor pattern. According to the characteristics of the self-built dataset, an improved scheme is proposed from the aspects of data processing and classifier selection. The mAP index of the improved C3D-SVM structure is improved by 3.1%.

(3) Design and implementation of comparative experiments. In this paper, a number of comparative experiments are carried out from the perspectives of network parameters, network structure, and dataset. It is found that using C3D-SVM behavior recognition model training enhanced FAD_Y dataset can obtain the best performance of behavior

classification. The average precision of the model is 0.9198, the average recall is 0.9091, and the F1 score is 0.9144.

Key words: dataset construction, behavior recognition, deep learning, C3D-SVM