DEMO: Human Activity Analysis Using Pose Recognition

Egor Rybin, Nikita Bazhenov, Dmitry Korzun
Petrozavodsk State University (PetrSU)
Petrozavodsk, Russia
{rybin, bazhenov, dkorzun}@cs.karelia.ru

***Abstract*—The demo is part of a system for multi-parameter monitoring of production machinery (technical state, operating conditions, personnel actions). In this demo, we develop a video service for human activity recognition for monitoring working personnel. The service implements neural networks to recognize the pose as skeleton and its movement.**

Tracking the actions of employees at industrial monitoring enterprises is an important and urgent task in our time. Unlike conventional human silhouette recognition, posture recognition allows you to determine the position of individual body parts in space.

The basic task of recognizing person's activity can be considered the recognition of his skeleton. The interpretation of the results obtained can be used to determine the interaction of the operator with other objects in the enterprise (people or machinery) to monitor personnel and provide timely assistance in the event of unforeseen circumstances.

The human activity recognition, in turn, has a number of problems. The problem of filtering out false positives among several suggested poses is assessed in [1]. The authors propose an approach aimed at the problem of imprecise values of the validity of the CNN application, which is also used for the final assessment of a person's posture.

Implementation of human activity recognition can be used in different video surveillance applications, e.g., monitoring what employees doing inside the factory territory, checking space near dangerous machines for a safety violation, examine outside for actions from potential intruders.

Human pose analysis is needed to find the person's skeleton to correlate actions with the pose. The method used in this work labels actions based on differences between seen and target (for a particular action) angles between limbs.

There exist more general methods [2]. They tuned to evaluate human actions directly from the images. Unfortunately, such methods would require a lot of images to learn a new pose, while the method used in this work will need only the simple description with new angles, though it may not be as accurate or even work with complex actions.

The simplicity of adding new poses to the classifier opens possibilities to create tailored solutions for non-ordinary actions, specific to particular deployment, without needs to capture a large dataset.

For an example, if you bend your knees 90 degrees and your hip also 90 degrees to your body you are most likely sitting. Fig. 1 shows the service recognizing that man in the image is sitting with a probability of 0.83 and standing with 0.16.

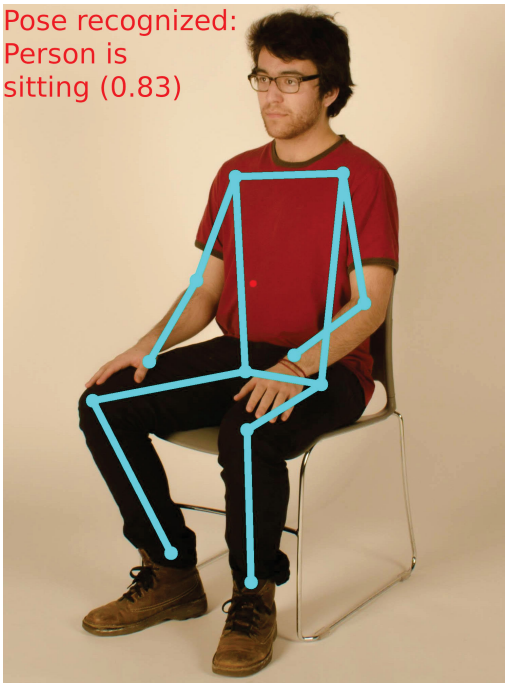

Fig. 1. A man is sitting on the chair.

Our demo implementation uses neural network PoseNet [3], developed by Google, to find and map a person on an image. Then method that corresponds mapped person joints to the particular action used. Moreover, this demo shows how the system works with single images and/or video streams. The high-level service architecture is shown in Fig. 2.

At first, a given image (or frame from a video stream) come into the neural network. The network then returns multiple probabilities of body parts. After postprocessing, or interpretation, of this results system will have list of joints with discreet coordinates. In the next step, the joints would be combined into body parts (e.g., hand, elbow and shoulder form arm together) and features based on their relative position would be counted. As the final step these features would be used to calculate what action the person performs.

ISSN 2305-7254

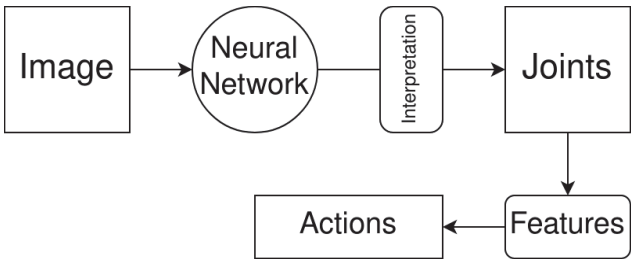

Fig. 2. Architecture of the system

ACKNOWLEDGMENT

This research is implemented in Petrozavodsk State University (PetrSU) with financial support by the Ministry of Science and Higher Education of Russia within Agreement no. 075-11-2019-088 of 20.12.2019 on the topic “Creating the high-tech production of mobile microprocessor computing modules based on SiP and PoP technology for smart data collection, mining, and interaction with surrounding sources”.

REFERENCES

[1] K. Gupta, L. Petersson, and R. Hartley, “Cullnet: Calibrated and pose aware confidence scores for object pose estimation,” in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, Oct 2019, pp. 2758–2766.

[2] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, “2d human pose estimation: New benchmark and state of the art analysis,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2014.

[3] G. Papandreou, T. Zhu, N. Kanazawa, A. Toshev, J. Tompson, C. Bregler, and K. Murphy, “Towards accurate multi-person pose estimation in the wild,” 2017.