Depth-Based Activity Recognition with Convolutional Neural Networks

Alisha Rege, Sanyam Mehra, Alyssa Vann
{amr6114, sanyam, avann}@stanford.edu

Use computer vision techniques to build an integrated solution for the remote monitoring, assessment, and support of seniors living independently at home.

Labelled data from depth sensors in the On Lok Senior Home in San Francisco

Build two state-of-the-art convolutional neural networks

Created a data labelling tool

9 depth sensors/ 4 rooms

12K video clips collected

28 Activities

Per frame labels for person, bounding box, activity

Video clips 7 x 224 x 224 x 1 (frame x width x height x channel)

Batch size of 4 video clips

Depth-Based Activity Recognition with Convolutional Neural Networks

Alisha Rege, Sanyam Mehra, Alyssa Vann
{amr6114, sanyam, avann}@stanford.edu

Use computer vision techniques to build an integrated solution for the remote monitoring, assessment, and support of seniors living independently at home.

Labelled data from depth sensors in the On Lok Senior Home in San Francisco

Build two state-of-the-art convolutional neural networks

Created a data labelling tool

9 depth sensors/ 4 rooms

12K video clips collected

28 Activities

Per frame labels for person, bounding box, activity

Video clips 7 x 224 x 224 x 1 (frame x width x height x channel)

Batch size of 4 video clips

Dataset

3D Convolutional Neural Network

- 3D Convolutional Layers
- Batch Normalization
- ReLU activations
- 3D Max Pooling
- 3D Average Pooling
- Fully Connected Layers

3D Convolutional Neural Network

- Pretrained Resnet for spatial feature extraction
- LSTM recurrent network for temporal structure

Recurrent Convolutional Neural Network

- Record and label more data
- Extend to person localization – from annotations created in dataset
- Activity recognition from point cloud representation
- Feature extraction using optical flow, depth-based descriptors
- Extend to other modalities; Thermal sensor data

Future Work

References