Interactive Gesture Plugin using Quantized Neural Network

Shaofan Lai
Gesture Recognition

- **Traditional methods: (Filter+Threshold)**
  - Sensitive to the background.
  - Hard-coded definition of gesture

- **State-of-the-art Devices: (Kinect, Leap Motion)**
  - Requires extra information (e.g. depth, infrared)
  - Expensive

- **Deep Neural Network (Lab setting):**
  - Intensive computation
  - Works with high-resolution video
  - Millions of data samples
Recognize Gesture with JeVois (OpenCV)

Code modified from: github.com/vipul-sharma20/gesture-opencv
Quantized (Binarized) Neural Network

Fig. 1: We propose two efficient variations of convolutional neural networks. **Binary-Weight-Networks**, when the weight filters contain binary values. **XNOR-Networks**, when both weight and input have binary values. These networks are very efficient in terms of memory and computation, while being very accurate in natural image classification. This offers the possibility of using accurate vision techniques in portable devices with limited resources.

More details can be found on: https://shaofanlai.com/post/47
XNOR-Network with Darknet

- On CIFAR-10
  - Standard Network
    - top 1: 0.831600, top 5: 0.992000
  - Xnor Network
    - top 1: 0.684600, top 5: 0.976100

- Drawback
  - No time advantage
    - GPU (GTX 1080 Ti)
      - Standard~2.872s vs Bin~5.340s
    - CPU (Intel 7700K)
      - Standard~32.908s vs Bin~33.316s
  - No memory/storage advantage
My Proposal

Using quantized neural network to recognize the gestures in my workplace:

- Using only JeVois Camera(s)
- Relatively consistent background
- Customized gestures
- Interacting with the system in real time
Challenge

- Training and employment of the gesture model
  - Data (Collected in workplace? Transferred from other gesture datasets?)
  - Model (Attention-base? Detection-based? Tracking-based?)
  - Training (Unsupervised (with saliency)? Supervised (with labels)? Reinforcement Learning (with auxiliary sensors)?)
- Implementation of the Quantized Neural Network
  - Memory usage improvement
  - Computation acceleration
- Hacking GNOME system plugin
  - Communication with JeVois
  - User Interface design
References

- Binarized Neural Networks: Training Neural Networks with Weights and Activations Constrained to +1 or -1
- Quantized Neural Networks: Training Neural Networks with Low Precision Weights and Activations
- XNOR-Net: ImageNet Classification Using Binary Convolutional Neural Networks