Face anti-spoofing CNN

July 7, 2016

1 Learn Convolutional Neural Network for Face Anti-Spoofing

1.1 Intro

Due to the diversity of spoofing attacks, existing face antispoofing approaches can be mainly categorized into four groups:

- Texture-based Anti-Spoofing: LBP, Difference of Gausssian, analysis of Fourier spectra.
- Motion-based Anti-Spoofing: eye blinking, lip movement classification and lip-reading, optical flow.
- 3D Shape-based Anti-Spoofing: 3D projective invariants.
- Multi-Spectral Reflectance-based Anti-Spoofing: utilize the illuminations beyond visual spectrum, reflection intensities, gradient-based multi-spectral,

1.2 Method

Data Preparation First, face is located with Viola-Jones algorithm, them is aliagned with a algorithm from another paper. Then, authors prepare the input images with five scales. Authors propose augment data temporally, they feed the CNN with more than one frame.

Feature Learning Authors implement a canonical CNN structure. They use the CNN which won ImageNet large scale visual recognition in 2012 which has five convolutional layers followed by thre fully conencted layers. Response normalizacion layers are used for the outputs of the first and second conv layers; Max-pool layers are used at the output of the first, second and last conv layers; the ReLu nnon-linearity is applied at the output of every conv layer and in the first two fully connected layers.

Clasification Use supervector machine with RBF Kernel to classify

1.3 Experiments

Image Settings Re-scale images with ratios 1.4, 1,8, 2.2, 2.6 All input images are resize to 128 x 128.

CNN Settings Authors use Caffe to bluid the Net. The earning rate is 0.001, the decay rate is 0,001 and the momentud during the raining is 0.9.

2 Learning Temporal Features Using LSTM-CNN Architecture for FaceAnti-spoofing

2.1 Intro

RNN architecture (recurrent neural networks) may suffer from optimization problem of exponential decay of gradient information. Thus, we implement a recurrent neural network with Long Short Term Memory (LSTM) units. The LSTM units can discover long-range temporal relationships from the input sequences by using input gates, output gates, forget gates to control modifying, accessing and storing the internal states. We put the LSTM layer above a convolutional neural network (CNN) architecture.

2.2 Method

Authos treat face anti-spoofing as video classification problem. he input of our model is the sequence of video frames (x 1 , x 2 , ..., x n), and the output is a binary number y indicating whether the input sequences are real.

2.2.1 LSTM

2.2.2 LSTM-CNN Architecture

The CNN has two convolutional layer and a max pooling layer after each, one fully connected layer, one dropout later and one softmax layer to predict.

The use of the dropout layer van significantly prevent over-fitting.

To learn temporal structures, authors put a LSTM layer between the fully connected and the softmax layer.

2.3 Training

Authos use caffe. Stochastic Gradient Descent is used with a momentum of 0.9. The learning rate is 0.001 for severals iterations.

2.4 Model detailss

CNN has two conv layers, the first one has 48 filters amd the second one 96, The size of the filters are 3x3 and the stride of 1 pixel.

The size of the pooling layer is 2 pixels and stride of 2 pixels.

The fully connected layer has 1000 neurons and each activation is set to 0 with a propabilbity of 0.5 during training.

The non-linear function is rectified linear units (ReLu).

The LSTM has 30 internal cells for each time steps.

3