Cascade Classifier Training Example

Version: OpenCV 3.4.0
Cascade Classifiers: https://en.wikipedia.org/wiki/Cascading_classifiers
Ensemble Learning: https://en.wikipedia.org/wiki/Ensemble_learning
Cascade Classifier Training -> **Ensemble Decision Tree Training (stage i based on stage i-1)**

Cascading is a particular case of **ensemble learning** based on the concatenation of several Classifiers, using all information collected from the **output from a given classifier as additional information for the next classifier in the cascade**. Unlike voting or stacking ensembles, which are multi-expert systems, cascading is a multistage one.

Cascading Classifiers are trained with several hundred "positive" sample views of a particular object and arbitrary "negative" images of the same size. After the classifier is trained it can be applied to a region of an image and detect the object in question. To search for the object in the entire frame, the search window can be moved across the image and check every location for the classifier. This process is most commonly used in image processing for object detection and tracking, primarily facial detection and recognition.

1. Prepare description file for positive images:

```
find positives/* -exec identify -format '%i 1 0 0 %w %h\n' {} \; > description.dat
```

2. Create positive samples:

```
opencv_createsamples -info description.dat -vec output.vec
```

```
Info file name: description.dat
Img file name: (NULL)
Vec file name: output.vec
BG  file name: (NULL)
Num: 1000
BG color: 0
BG threshold: 80
Invert: FALSE
Max intensity deviation: 40
Max x angle: 1.1
Max y angle: 1.1
Max z angle: 0.5
Show samples: FALSE
Width: 24
Height: 24
Max Scale: -1
```

```
RNG Seed: 12345
Create training samples from images collection...
description.dat(83) : parse errorDone. Created 82 samples
```

Since we didn't specify the number of samples to generate, it will throw parse error when all 82 positive images are used up. To remove the error in the end, you could specify the number of samples by adding flag `-num`.

3. Train cascade classifier:

```
opencv_traincascade -vec output.vec -data cascade_dir/ -bg negatives.txt
-numPos 66 -numNeg 4000
PARAMETERS:
cascadeDirName: cascade_dir/
vecFileName: output.vec
bgFileName: negatives.txt
numPos: 66
numNeg: 4000
numStages: 20
precalcValBufSize[Mb] : 1024
precalcIdxBufSize[Mb] : 1024
acceptanceRatioBreakValue : -1
stageType: BOOST
featureType: HAAR
sampleWidth: 24
sampleHeight: 24
boostType: GAB
minHitRate: 0.995
maxFalseAlarmRate: 0.5
weightTrimRate: 0.95
maxDepth: 1
maxWeakCount: 100
mode: BASIC
Number of unique features given windowSize [24,24] : 162336

===== TRAINING 0-stage =====
<BEGIN
POS count : consumed   66 : 66
NEG count : acceptanceRatio    4000 : 1
Precalculation time: 29
+----+---------+---------+
|  N |    HR   |    FA   |
+----+---------+---------+
```

```
|  1|          1|      1|
+---+---------+---------+
|  2|          1|      1|
+---+---------+---------+
|  3|          1|  0.0655|
+---+---------+---------+
END>
Training until now has taken 0 days 0 hours 1 minutes 21 seconds.

===== TRAINING 1-stage =====
<BEGIN
POS count : consumed   66 : 66
NEG count : acceptanceRatio    4000 : 0.118811
Precalculation time: 31
+---+---------+---------+
|  N |   HR    | FA    |
+---+---------+---------+
|  1|          1|      1|
+---+---------+---------+
|  2|          1|      1|
+---+---------+---------+
|  3|          1|      1|
+---+---------+---------+
|  4|          1|  0.20225|
+---+---------+---------+
END>
Training until now has taken 0 days 0 hours 3 minutes 5 seconds.

===== TRAINING 2-stage =====
<BEGIN
POS count : consumed   66 : 66
NEG count : acceptanceRatio    4000 : 0.0162725
Precalculation time: 30
+---+---------+---------+
|  N |   HR    | FA    |
+---+---------+---------+
|  1|          1|      1|
+---+---------+---------+
|  2|          1|      1|
+---+---------+---------+
|  3|          1|      1|
+---+---------+---------+
|  4|          1|  0.1735|
```

```
+----+---------+---------+
END>
Training until now has taken 0 days 0 hours 4 minutes 52 seconds.

===== TRAINING 3-stage =====
<BEGIN
POS count : consumed   66 : 66
NEG count : acceptanceRatio      4000 : 0.00298203
Precalculation time: 31
+----+---------+---------+
|  N |   HR    |  FA     |
+----+---------+---------+
|   1|        1|        1|
+----+---------+---------+
|   2|        1|        1|
+----+---------+---------+
|   3|        1|        1|
+----+---------+---------+
|   4|        1|     0.24|
+----+---------+---------+
END>
Training until now has taken 0 days 0 hours 6 minutes 40 seconds.

===== TRAINING 4-stage =====
<BEGIN
POS count : consumed   66 : 66
NEG count : acceptanceRatio      4000 : 0.000810957
Precalculation time: 31
+----+---------+---------+
|  N |   HR    |  FA     |
+----+---------+---------+
|   1|        1|        1|
+----+---------+---------+
|   2|        1|        1|
+----+---------+---------+
|   3|        1|     0.18|
+----+---------+---------+
END>
Training until now has taken 0 days 0 hours 8 minutes 19 seconds.

===== TRAINING 5-stage =====
<BEGIN
POS count : consumed   66 : 66
```

```
NEG count : acceptanceRatio      4000 : 0.000157721
Precalculation time: 29
+----+---------+---------+
|  N |   HR    |  FA     |
+----+---------+---------+
|   1|        1|        1|
+----+---------+---------+
|   2|        1|        1|
+----+---------+---------+
|   3|        1|        1|
+----+---------+---------+
|   4|        1|  0.33025|
+----+---------+---------+
END>
Training until now has taken 0 days 0 hours 11 minutes 16 seconds.

===== TRAINING 6-stage =====
<BEGIN
POS count : consumed   66 : 66
NEG count : acceptanceRatio      4000 : 4.98294e-053903
Precalculation time: 31
+----+---------+---------+
|  N |   HR    |  FA     |
+----+---------+---------+
|   1|        1|        1|
+----+---------+---------+
|   2|        1|        1|
+----+---------+---------+
|   3|        1|        1|
+----+---------+---------+
|   4|        1|        1|
+----+---------+---------+
|   5|        1|  0.29675|
+----+---------+---------+
END>
Training until now has taken 0 days 0 hours 17 minutes 22 seconds.

===== TRAINING 7-stage =====
<BEGIN
POS count : consumed   66 : 66
NEG count : acceptanceRatio      4000 : 1.63696e-05
Precalculation time: 29
+----+---------+---------+
```

```
|  N |   HR   | FA   |
+----+--------+------+
|   1|       1|     1|
+----+--------+------+
|   2|       1|     1|
+----+--------+------+
|   3|       1|     1|
+----+--------+------+
|   4|       1|  0.354|
+----+--------+------+
END>
Training until now has taken 0 days 0 hours 31 minutes 19 seconds.

===== TRAINING 8-stage =====
<BEGIN
POS count : consumed   66 : 66
NEG count : acceptanceRatio    4000 : 5.96971e-06
Precalculation time: 29
+----+--------+------+
|  N |   HR   | FA   |
+----+--------+------+
|   1|       1|     1|
+----+--------+------+
|   2|       1|     1|
+----+--------+------+
|   3|       1|     1|
+----+--------+------+
|   4|       1|  0.19175|
+----+--------+------+
END>
Training until now has taken 0 days 1 hours 7 minutes 20 seconds.

===== TRAINING 9-stage =====
<BEGIN
POS count : consumed   66 : 66
NEG count : acceptanceRatio    0 : 0
Required leaf false alarm rate achieved. Branch training terminated.
```

It terminated at stage 9 because the default minimum hit rate and maximum false alarm rate has been satisfied. So far, we get the **cascade.xml** in the directory we specified in the command.