Suggested steps for creating a cascade detector

The OpenCV documentation for these steps can be found in:

```
http://docs.opencv.org/doc/user_guide/ug_traincascade.html
```

The description here is mostly from the following link:

```
http://www.memememememememe.me/training-haar-cascades
```

1. Use **imageclipper** to create the images to be used as positive samples. You may use instead any other software, for example photoshop or Microsoft paint. Keep them in a subfolder called **pos**.

2. Create a subfolder called **neg** and put in it the negative samples. The image size of the negative samples should be bigger than the image size of the positive samples.

3. Create the file **neg.txt** with a text editor. There should be a line for each negative sample. Thus, if **fruits .jpg** is one of the negative samples, neg.txt should have the line: **neg/fruits .jpg**.

   Because of apparent bug in **createsamples** please make sure that there are as many lines in **neg.txt** as the number of examples you want to generate from a single positive example. For example, in order to generate 100 examples from each positive example, make sure that **neg.txt** contains at least 100 lines. This can be done by duplicating the lines as many times as needed.

4. For each positive sample run **createsamples.exe** to create many images. E.g:

   ```
   createsamples.exe -img pos\p1.png  -bg neg.txt
           (continue same line) -info crop1.txt -num 100 -w 20 -h 20
   createsamples.exe -img pos\p2.png  -bg neg.txt
           (continue same line) -info crop2.txt -num 100 -w 20 -h 20
   ```

   At this point you should see the generated images in the current folder. The files **crop1.txt** and **crop2.txt** contain a list of these images.

5. Create one big file by concatenating all the info files. This can be one on windows by the following command:

   ```
   copy /b crop1.txt + /b crop2.txt crop.txt
   ```

6. Create the **.vec** file.

   ```
   createsamples.exe -info crop.txt -bg neg.txt -vec cropped.vec
           (continue same line) -num 200 -w 20 -h 20
   ```

   The number of samples generated here should be no bigger than the total number of samples available. You may encounter a bug that can be avoided by reducing that number.

   The values of w and h (width and height) are very important for the run time. Make sure they are not too big. They should also be identical in all runs.

7. **Train.**

   ```
   traincascade.exe -data . -vec cropped.vec -bg neg.txt
     (continue same line) -w 20 -h 20 -numPos 200 -numStages 13
   ```

   This creates **cascade.xml**.

8. **Test.**

   ```
   Detect.exe cascade.xml  ..\..\images\fingers
   Detect.exe cascade.xml
   ```

Notes

- **traincascade.exe** may run for a long time. You can terminate it before it is done and run it again with a smaller number of stages. In this case it remembers the stages already calculated in the previously terminated run.
- **traincascade.exe** always remembers the previous run by creating the files **stagei.xml, i=0,1,..** and **params.xml**. You should remove these files before running a new experiment.