

—

Blocking the Nude Image in Advertisement at Runtime

Introduction



Skin colour detection has been used in numerous computer vision applications like face detection, nudity recognition, hand gesture detection and person identification. Skin colour detection is a challenging task as the skin colour in an image is sensitive to various factors like illumination, camera characteristics, ethnicity, individual characteristics such as age, sex and body parts and other factors like makeup, hairstyle and glasses. All these factors affect appearance of skin colour. Another problem is that there is a significant overlap between the skin and non-skin pixels. However when these techniques are used in real-time, it is crucial to follow time deadlines and memory constraints. Sometimes, accuracy may need to be sacrificed when the skin detection strategy is used only as a preprocessing step to face detection, particularly in real time applications. In this study we have focused on the problem of developing an accurate and robust model for the human skin.

Objective



The model will offer child proof surfing on the internet without parent intervention. So that parents do not have to worry about their children coming across nude images at such an early age without knowing the actual meaning of it. That is it will monitor every page and filter all the images, hiding their details from the children and disabling their activation upon any click. Even if the people in the image is not completely nude, may it be just the upper or lower half of the person's body; the model will still blur that image once it reaches the minimum percentage of human pixel colour set by the classifier, thus ensuring guaranteed protection.

Face Detection



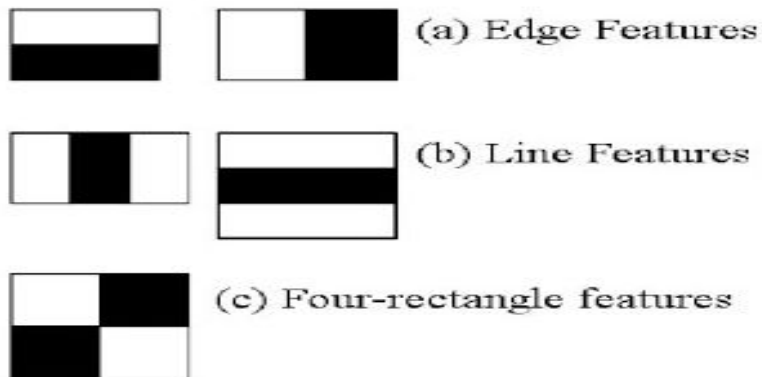
Face Detection is a type of application classified under “computer vision” technology. It is the process in which algorithms are developed and trained to properly locate faces or objects in images. These can be in real time from a video camera or from photographs. On social media apps like Snapchat, face detection is required to augment reality which allows users to virtually wear dog face masks using fancy filters. Another use of face detection is in smartphone face ID security.

So the first aim of our project is to detect whether there is a human face in the image or not. Face Detection is implemented in OpenCv using 2 techniques:

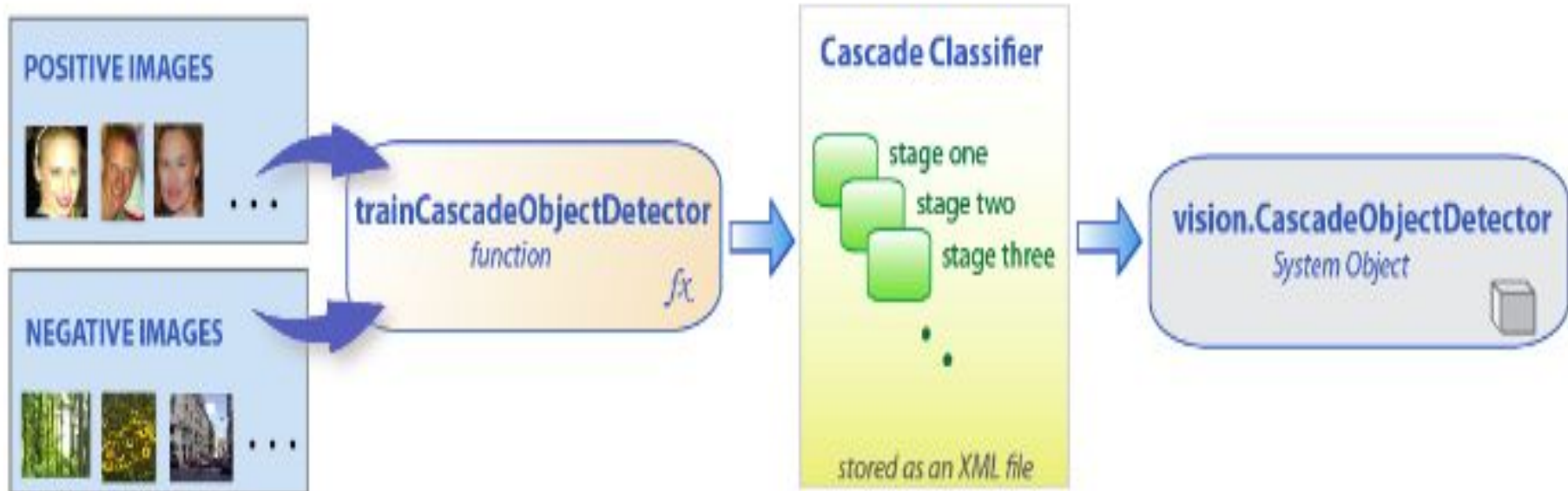
- 1) Haar Cascade Classifier
- 2) Deep Neural Network

Haar Cascade Classifier

Object Detection using Haar feature-based cascade classifiers is an effective object detection method proposed by Paul Viola and Michael Jones. It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images. Initially, the algorithm needs a lot of positive images (images of faces) and negative images (images without faces) to train the classifier. Then we need to extract features from it. They are just like our convolutional kernel. Each feature is a single value obtained by subtracting sum of pixels under white rectangle from sum of pixels under black rectangle.



Haar Cascade Classifier



Deep Neural Network



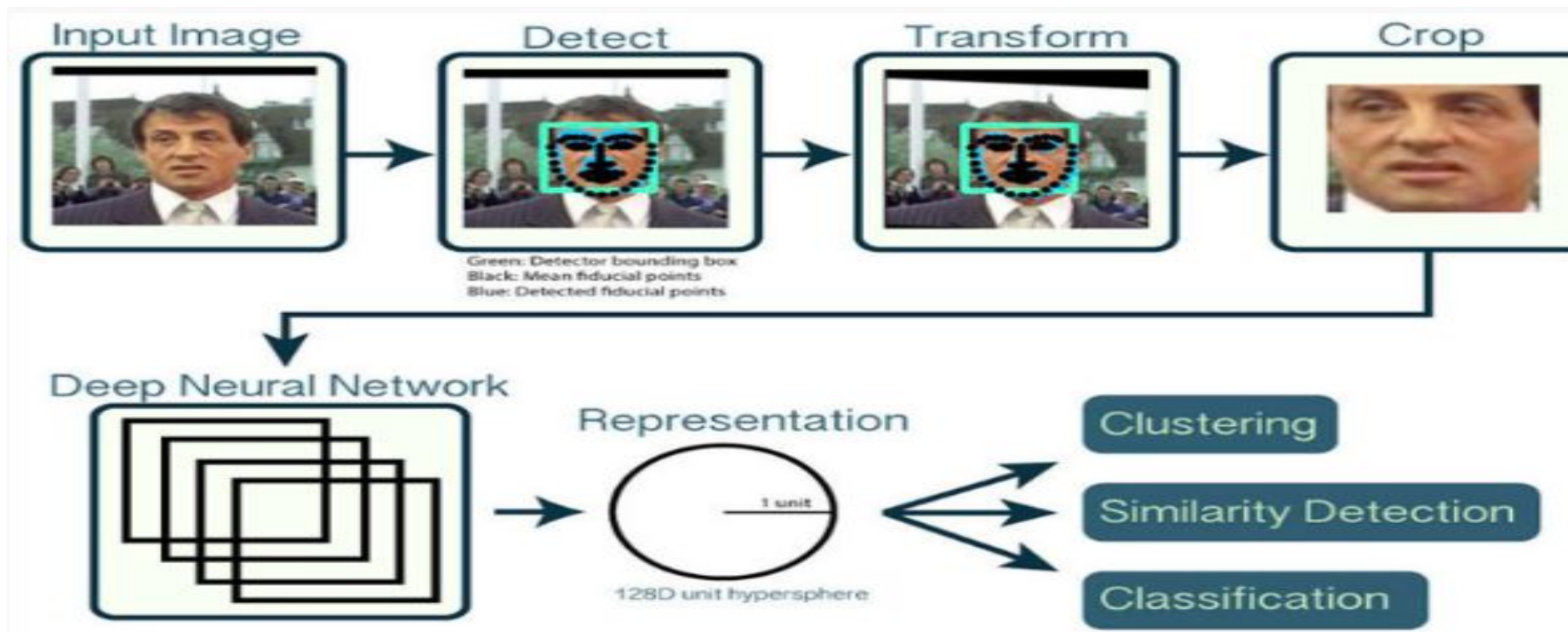
OpenCV (Open Source Computer Vision) is a library with functions that mainly aiming real-time computer vision. OpenCV supports Deep Learning frameworks Caffe, Tensorflow, Torch/PyTorch. With OpenCV we can perform face detection using pre-trained deep learning face detector model which is shipped with the library. When OpenCV 3.3 was officially released, it has highly improved deep neural networks (dnn) module.

We have used Caffe based framework model for face detection. To use OpenCV Deep Neural Network module with Caffe models we will need two files .

- 1) prototxt file which defines model architecture
- 2) caffemodel file which contains the weights for the actual layers

OpenCV's face detector is based on the Single Shot Detector framework with a ResNet base network.

Deep Neural Network



Neural Networks



Neural Networks is a machine learning algorithm, which is built on the principle of the organization and functioning of biological neural networks. Neural networks consist of individual units called neurons. Neurons are located in a series of groups — layers. Neurons in each layer are connected to neurons of the next layer. Data comes from the input layer to the output layer along these compounds. Each individual node performs a simple mathematical calculation. Then it transmits its data to all the nodes it is connected to.

Different types of neural network architecture are as follows:

- 1) Perceptrons
- 2) Convolutional Neural Networks
- 3) Recurrent Neural Networks
- 4) Long / Short Term Memory
- 5) Gated Recurrent Unit
- 6) Hopfield Network

Why Convolution Neural Networks

Convolutional neural networks (CNN) is a special architecture of artificial neural networks, proposed by Yann LeCun in 1988. CNN uses some features of the visual cortex. One of the most popular uses of this architecture is image classification. For example Facebook uses CNN for automatic tagging algorithms, Amazon – for generating product recommendations and Google – for search through among users' photos. The main task of image classification is acceptance of the input image and the following definition of its class. This is a skill that people learn from their birth and are able to easily determine that the image in the picture is an elephant. But the computer sees the pictures quite differently:

What I see

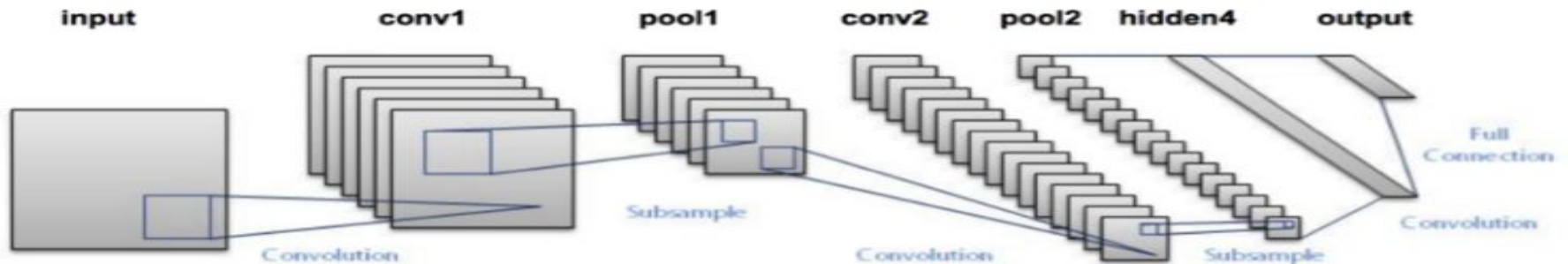


What a computer sees


08	02	22	97	38	15	00	40	00	75	04	05	07	78	52	12	50	77	91	08
49	49	99	40	17	81	18	57	60	87	17	40	98	43	69	48	04	56	62	00
81	49	31	73	55	79	14	29	93	71	40	67	53	88	30	03	49	13	36	65
52	70	95	23	04	60	11	42	69	24	68	56	01	32	56	71	37	02	36	91
22	31	16	71	51	67	43	89	41	92	36	54	22	40	40	28	66	33	13	80
24	47	32	60	99	03	45	02	44	75	33	53	78	36	84	20	35	17	12	50
32	98	81	28	64	23	67	10	26	38	40	67	59	54	70	66	18	38	64	70
67	26	20	68	02	62	12	20	95	63	94	39	63	08	40	91	66	49	94	21
24	55	58	05	66	73	99	26	97	17	78	78	96	83	14	88	34	89	63	72
21	36	23	09	75	00	76	44	20	45	35	14	00	41	33	97	34	31	33	95
78	17	53	28	22	75	31	67	15	94	03	80	04	62	16	14	09	53	56	92
16	39	05	42	96	35	31	47	55	58	88	24	00	17	54	24	36	29	85	57
86	56	00	48	35	71	89	07	05	44	44	37	44	60	21	58	51	54	17	58
19	80	81	68	05	94	47	69	28	73	92	13	86	52	17	77	04	89	55	40
04	52	08	83	97	35	99	16	07	97	57	32	16	26	26	79	33	27	98	66
88	36	68	87	57	62	20	72	03	46	33	67	46	55	12	32	63	93	53	69
04	42	16	73	38	25	39	11	24	94	72	18	08	46	29	32	40	62	76	36
20	69	36	41	72	30	23	88	34	62	99	69	82	67	59	85	74	04	36	16
20	73	35	29	78	31	90	01	74	31	49	71	48	86	81	16	23	57	05	54
01	70	54	71	83	51	54	69	16	92	33	48	41	43	52	01	89	19	47	48

Convolution Neural Network cont.....

Instead of the image, the computer sees an array of pixels. For example, if image size is 300 x 300. In this case, the size of the array will be 300x300x3. Where 300 is width, next 300 is height and 3 is RGB channel values. The computer is assigned a value from 0 to 255 to each of these numbers. This value describes the intensity of the pixel at each point. To solve this problem the computer looks for the characteristics of the base level. In human understanding such characteristics are for example the trunk or large ears. For the computer, these characteristics are boundaries or curvatures. And then through the groups of convolutional layers the computer constructs more abstract concepts. In more detail: the image is passed through a series of convolutional, nonlinear, pooling layers and fully connected layers, and then generates the output.



Output of Cat and Dog Tutorial with 4 Layers



```
Epoch 11/20
32/32 [=====] - 229s 7s/step - loss: 0.6069 - acc: 0.6576 - val_loss: 0.6123 - val_acc: 0.6475
Epoch 12/20
32/32 [=====] - 228s 7s/step - loss: 0.6024 - acc: 0.6834 - val_loss: 0.6115 - val_acc: 0.6625
Epoch 13/20
32/32 [=====] - 227s 7s/step - loss: 0.6050 - acc: 0.6580 - val_loss: 0.6072 - val_acc: 0.6600
Epoch 14/20
32/32 [=====] - 226s 7s/step - loss: 0.6050 - acc: 0.6761 - val_loss: 0.5730 - val_acc: 0.7000
Epoch 15/20
32/32 [=====] - 230s 7s/step - loss: 0.6051 - acc: 0.6700 - val_loss: 0.6301 - val_acc: 0.6250
Epoch 16/20
32/32 [=====] - 230s 7s/step - loss: 0.5850 - acc: 0.6917 - val_loss: 0.6203 - val_acc: 0.6450
Epoch 17/20
32/32 [=====] - 227s 7s/step - loss: 0.6045 - acc: 0.6427 - val_loss: 0.5674 - val_acc: 0.7250
Epoch 18/20
32/32 [=====] - 230s 7s/step - loss: 0.5942 - acc: 0.6790 - val_loss: 0.5915 - val_acc: 0.6575
Epoch 19/20
32/32 [=====] - 230s 7s/step - loss: 0.5915 - acc: 0.6861 - val_loss: 0.5500 - val_acc: 0.7575
Epoch 20/20
32/32 [=====] - 225s 7s/step - loss: 0.5653 - acc: 0.7142 - val_loss: 0.5567 - val_acc: 0.7275
>72.750
```

Using VGG16 model



```
Epoch 1/10
32/32 [=====] - 1192s 37s/step - loss: 0.9444 - acc: 0.9191 - val_loss: 0.5645 - val_acc: 0.9500
Epoch 2/10
32/32 [=====] - 1181s 37s/step - loss: 0.3874 - acc: 0.9700 - val_loss: 0.3456 - val_acc: 0.9750
Epoch 3/10
32/32 [=====] - 1187s 37s/step - loss: 0.3833 - acc: 0.9719 - val_loss: 0.6194 - val_acc: 0.9525
Epoch 4/10
32/32 [=====] - 1192s 37s/step - loss: 0.3094 - acc: 0.9751 - val_loss: 0.5087 - val_acc: 0.9600
Epoch 5/10
32/32 [=====] - 1185s 37s/step - loss: 0.1624 - acc: 0.9858 - val_loss: 0.2661 - val_acc: 0.9750
Epoch 6/10
32/32 [=====] - 1189s 37s/step - loss: 0.1257 - acc: 0.9902 - val_loss: 0.2903 - val_acc: 0.9750
Epoch 7/10
32/32 [=====] - 1200s 37s/step - loss: 0.1794 - acc: 0.9873 - val_loss: 0.6117 - val_acc: 0.9525
Epoch 8/10
32/32 [=====] - 1205s 38s/step - loss: 0.1362 - acc: 0.9902 - val_loss: 0.4027 - val_acc: 0.9725
Epoch 9/10
32/32 [=====] - 1207s 38s/step - loss: 0.1283 - acc: 0.9902 - val_loss: 0.2983 - val_acc: 0.9800
Epoch 10/10
32/32 [=====] - 1190s 37s/step - loss: 0.1441 - acc: 0.9897 - val_loss: 0.3136 - val_acc: 0.9800
>98.000
```

Predictions



```
/content/drive/My Drive/major projects/cat and dog/test/dog.4024.jpg  
[1.]  
/content/drive/My Drive/major projects/cat and dog/test/cat.4134.jpg  
[0.]  
/content/drive/My Drive/major projects/cat and dog/test/dog.4025.jpg  
[1.]  
/content/drive/My Drive/major projects/cat and dog/test/cat.4183.jpg  
[0.]  
/content/drive/My Drive/major projects/cat and dog/test/dog.4166.jpg  
[1.]  
/content/drive/My Drive/major projects/cat and dog/test/cat.4138.jpg  
[0.]  
/content/drive/My Drive/major projects/cat and dog/test/cat.4040.jpg  
[0.]  
/content/drive/My Drive/major projects/cat and dog/test/cat.4162.jpg  
[0.]  
/content/drive/My Drive/major projects/cat and dog/test/dog.4102.jpg  
[1.]  
/content/drive/My Drive/major projects/cat and dog/test/dog.4134.jpg  
[1.]  
/content/drive/My Drive/major projects/cat and dog/test/cat.4014.jpg  
[0.]  
/content/drive/My Drive/major projects/cat and dog/test/dog.4014.jpg  
[1.]  
/content/drive/My Drive/major projects/cat and dog/test/dog.4033.jpg  
[1.]
```



Thank You