Simple Face Detection Method Using Convolutional MAtrix

Aida Ghayour-Khiavi
Department of Computer Sciences
Memorial University of Newfoundland
aghayourkhia@mun.ca

April 6, 2019

Abstract

Detecting human faces in a ditgital image is one of the most challenging fields in the area of Compuer Vision. Variant appearnces of indviduals, different facial characteristics, presence of some features along with wide range of poses one can obtain in the input image, requires variant approaches for the face detection purposes. A face detecting system is expected to locate all the faces present in a picture regardless of their pose or characteristics, and return them with some possible further information. We have purposed an Face Detection algorithm based on Machine learning with a focus on low computational cost. In this method, we move the original image to a lower resolution, after an image processing phase, some candidate regions are extracted followed by a verification phase in which machine learnign techniques are employed to determine if the region includes a face or not.

Keywords:*Computer vision,Face detection,*

I. Introduction

[36] Face detection is can refer to the problem of "given still or video images of a scene, identify one or more persons in the scene using a stored database of faces."[35]. In this matter, many challenges are faced as the patterns based on human characteristics, can vary widly.Some of the challenges are listed as following:

- **Pose**: Human faces can show up in variant poses in an image. The direction of their face and the oreitation they obtain related to camera, result in variant looks in an image which can be challenging for a system to classify them as same object (face).
- **Existance of structual components**: The general structure of human face can be affacted by the presence of some features such as glasses or beard. This puts the system in trouble of considering wise range of human faces, while keeping all the crierias as specific as possible to end up in an accurate result.
- **Facial Expressions**: All the actions we need to use our facial feaatures to take, such as smiling or crying, has a big effect on the way a face looks in an image.
- **Occulsion**: The border of faces present in an image are not specificly distinguishable al lthe time. Some faces can be over lapping or covered by some object in front of them which disturbs their pattern.
- **Image orientation**: It should be considered that not all the faces in an image are perfectly vertical. Camera, or person's orientation relative to camera can range widely.

To go further in this field, it is neccessary to be familiar with some metaphors. Although

there is no coherent defintion of face detection, but in here we use the term as finding all the faces in an arbitary inout image. This system can be considered as a more specific task than face localization which can be interpreted as the procedure of localizing a face in an image containing only one face. Also, it is essential to be familiar with the terms "False Negative" and "False Positive" as they are two factors on which our evaluation methods based. False negative refers to those paches of image which contain a face but has not been returned as one by the system. False Positives though, are the part of image which has been counted as a face but they are not. [36] Face detection system's evaluation are done by feeding some standard set of images and counting the number of False Positives and False Negatives. It should be mentioned that each system is focused on a specific set of problems, and aims to achieve a particular goal. Hence, each system is to be evaluated using suitable data base. In the following sections, we will go through some of most well-known implementations categorized into four major group. At the end, we will review some mostly used standard data bases.

II. Face Detection

It has already been mentioned that face detection area is very broad and challenging. Thus, it is hard to categorize all the implemented systems into independent groups, but generally it would be more appropriate to put them into classes based on the target of their system and the methods they employ to tackle the present obstacles. There are four major categories ar purposed in survey by Yang et al. [36] as following:

- **Knowledge-based methods**: These methods implement the general knowledge of human about faces and use those information mostly for face localization purposes. Generally, they employ the relative characteristics of human face's componants.
- **Feature invariant approaches**: Human face, regardless of its features which can be affacted greatly by enviromental factors such as lightening, has some specific structures, which can be used for face detection purposes.
- **Template matching methods**: It is obvious that all human faces and its particular segments, have an standard pattern. Templating matchinf mthods, detect faces by relying on correlation of these patterns and the input image.
- **Appearance-based methods**: In addition to the abovementioned methods, some systems tackle deal with this problem by learning face's variations from a set of tarining images.
  On the other hand, Mayank Chauhan et al [37] purposes some other categories such as Geometric based methods, High Level LAnguage Methods and Haar like Feature based methods.

In the followinf sections we are going to review some of purposed methods in each caegory.

i. Knowledge-Based Top-Down Methods

: These methods employ human's knowledge about human face and try to implement those as a system. Although it can be easy to come up with ideas about face's shape, encoding those ides are not an easy task to deal with. The difficulty raises from the fact that general rules can end up in many false positives and very specific ones may miss most of faces.
One of thses methodes were purposed by Yang and Huang.[38]. In their method, the input image unndergoes through three levels of proccessing with different resolutions. At first level, they look for candidate regions for a face. in the second layer, they apply histogram equalization and edge detection operators on the candidates returned from level two, on a higher resolution. Finally, on the third level they look for facial features in the candidated returned from second level. They purposed level deals

with cluttered back ground and scale variations.
Kotropoulos and Pilas[5], was motivated by the idea of mosaie image eployed by Yang and Huang and developed and idea to determine the bounderies of a face in an image by projecting it on horizentally and vertically followed by finding local minimas. The same method is applied on candidate face region to locate facial features thus validating them.Their method performed s satisfying job of 86.5 percent on images that were only containing one face in harmonious background.

ii. Bottom-Up Feature-Based Methods

Human being has no problem recognizing and identifying faces. This fact motivated scientiscts to delevop some models that are able to encode human's knowledge and use that information for face detection purposes, generally using edge based images. Location o and the relationship of the eyes, nostrils, mouth and eyebrows are widely employed in such methods. The most challenging obstcle in this case, is the feafures can be greatly affected by conditional factors such as illlumination thus not being easy to target.

ii.1 Facial Features

One of the interesting methods which is designed to detect all the faces larger than 60*60 pixel, in different poses and illuminations, is purposed by Yow and Cipolla [6]. In their method, they apply a Gaussian filter on the raw image to draw some interest points based on local mimimal of the result image. based on some factors, these interest point are grouped together to form a feature vector. A reashold is computed with the help of a training set and formation of covariance matrix. a candidate is validized as a face if Mahalanobic distance of the feature vecors were below a treshold.
A two phase method also is purposed by Amit et al.[7] for detecting frontal view of faces in still grey images with arbitary background. They tack the probelm in two levels: First, they try to focus on a particular region by spatial arrangement of edge fragments which is returned from a edge detection operator with the help of intensity difference. Then some spaitial arragments which are more likely to present in facial regions than in background, is selected among a larg number of training images. As the second level, all canididate regions are classified as either background or face using a learned CART tree.[8]

ii.2 Texture

Texture of human face's can become a handy tool in face detection area. One of the methods which is relying of human's face texture, is purposed by Dai and Nakano[9] which is based on second-order statistical features[10]. They ehance orange-like parts of input image thus employing face color in their designed method which makes it rubost against pose variations.

ii.3 Skin Color

Human skin color can be a truly promising tool in the field of human face detection. Skin patches in an image can be labeled usig variant color spaves such as RGB[11] or HSV[12]. Variant models have been purposed in this matter. THe simplest one is designed by Chai et al [13] in which they are lableing pixles as skin if they fall in a range which indicates skin color. Saxe and Foulds[14], purposed a method in which they ask user to select a pixel among skin pixels following wiht seeding it into their segmentation componant as control factor. In addition to these non-parametric algorithms, some metric methods are disigned to perform better in case of different skin colors, in a Gausian distirbued space. However, due to sensitivity of this parameter against different illuminations, skin color is not a very effective way to be employed as a face detection factor all by itself, although it is a very efficient method in case it is combined with other approaches.

ii.4 Multiple Features

Some methodes use some of features all together to deal with face detection problem. For example, skin color segmentation can be used to lable pixels, and then the result eill be explored for oval shaped patches reasulting in face detection. Normally these methods locate the faces and valid the candidate inputs with a help of feature extraction.[**?**]Sobottka and Pitas[15] purposed a model based on skin color and shape. In the first phase, they form patches of skin color pixels on HSV color space. Then the fittest ellipse is mapped on each regions following by verification phase based on feature extraction.In addition to methods which employ pixels to tackle the problem, some structeu,color and geometry based mthods are purposed. Yang and Ahuja[16] developed a method which caputures chromatic properties using statistical analysis. A multiscale segmentation us employed for forming homogeneous regions. Then regions are merged leading to an approximate shape of ellipse.Feature extraction methods then applied on these regions for validation.

iii. Template Matching

Template mataching is one of the simplest methos to be implemented. Such methods, predefine a standard face pattern (Frontal Faces), then for each input image, correlation values for facial features are computed. The validation relies on existance of these values. It should be noted that such methods are not efficient enough in ase of different scales and shapes. SOme mthods are purposed to deal with later problems more effectively.

iii.1 Predefined Templates

Sakai et al.[17] purposed a method in which they first focus on regions of interest. They extract candidate regions by the use ofnline segmentation, followed by a validation phase which sudies facial features in greater detail. The novel idea of region of interest is empolyed in later works. Craw et al.[18] presented a method in which they detect faces with the help of shape templates of edge images. After locating a candidate face, the same process in repeated on varian scales to extracr facial features. Govindaraju et al.[19][20][21] developed a two level method in which facial features are defined based on edges, then eomplyed to define a sides of a frontal face (Left,Right and hairdline).The main approach is implemented using a Marr-Hildreth edge operator and then linking the fragmented contours. COrnor detection is used for segmentation of these rehgions. Then these candidates are explored for facial features based on relative ratio of them.

iii.2 Deformable Templates

Cootes and Taylor[22] presented a method in which they select rectangular patches of image as candidate faces. If dfeatures fall in a range of above a treshold, they will undero varification phase with the help of Active Shape Model (ASM) search.

iv. Appearance-Based Methods

Appearance based methods work different from Template maching ones on Tamplate formation. These methods, normally rely on Machine learning or statistical methods to explore face or non-face regions with the help of examle images.Normally input images are reduced to a lower resolutuion to avoid expensive computations. One of the methods in this matter, is to define a discriminant function (as a tresholde) between a face and non-face candidates. For this purpose, normally image examples are reduced to a lower dimention to form the function (usually with the help of distance measurements).[23] Also, there are mthods in which Support Vector Machines and some other kernel based methods are employed. THere systems project the example to a higher dimension and then decide between a face and non-face pattern.

iv.1 Eigenfaces

Turk and Pentlandcite[23] presented a method based on principal component analysis. In their method, face and non-face training images are projected onto same spaced and clustered. The buttomline of this idea is the fact that non-face patterns vary exteremly from one image to another, unlike face examples. To varify the presence of a face, the distance from face space to the image resgion is calculated for all the locations in an image. The face is detected regarding local minima of the formed face map.

v. FACE IMAGE DATABASES AND PERFORMANCE EVALUATION

Most of the purposed mthods, need some standard data base to train their systems. On the other hand, it is esseintial for all the methods to ebe tested using a coherent and standard set of images especially for comparsion purposes. In this section, some of the most well-known data sets, slong with their main charactersitics are depicted.

v.1 Face Image Database

One of the mostly used data bases is FERET image set.[24] It is a collection of bust images of a oerson on a uniform background. It's main purpose is to develop face recognition systems. Turk and Pentland's data base[23] consists of frontal faces with head pose variations on a cluttered background. AT& T [25] is a set of 1o variant images for forty subjects, under different illumination with or without presence of facial expressions and features(Glasses) during a period of time. The Harvard database [26] inclused cropped and masked frontal faces under variant lightenings. Yale Face Database [27]contains 10 frontal image for each subject wih varian facial expressions, with or without glasses and under different lightenings. The Purdue AR dataset contains more than 3000 color images of both males and females un frontal views. Faces have variant expressions under varying lightening

v.2 Test Image Benchmarks

THe Face Image datasets are developed for face recognition purposes thus contain only one face per each image. There is a need for soe standard data bases to evaluate our system especially for simluationg real world problems also fro comparsion purposes.Sung and Poggio presented two sets of databases for face detection. [29][30]. First set is including frontal and close-frontal mugshot images with lighting variation. The second one, consists of images in such fashion that face only vovers a small region of image on a cluttered background. The most used face detection databse belongs to Rowlet ay al. [31][32] which is a set of almost 500 frontal faces in 130 images. Most of images contain a face on a cluttered backgrounf which makes it asuitable candidate for evaluation of frontal face detectors. In addition to these, some sets has been presented by Kodak,[33] which contains faces from low resulation to large ones.

III. Purposed Model

This purposed model includes 5 major phases: Image Processing, Corner Detection, Detection of Interest Region, Candidate Extraction and Verification. Each of these steps are precisely described in the following.

i. Image Processing

In this phase, eah image is processed and then reduced to a lower resolution to avoid expensive computational cost. In addition to that, this low resolution works in favor of the following procedure -Corner Detection- as it is simpler to detect corners combined by bigger and fewer pixels. Briefly, algorithm takes the oroginal image, turns it to a black and white one and with the help of some filters (Such as Guasian filter) edges of the image is enhanced in a binary space. This edge image then is resized leading to a lower resolution to be feed to the next phase. The original and binary edge image are shown in figures 1 and 2 respectievly. As it is illustrated, a colorful image of the size

Figure 1: *Original Image*

Figure 2: *Binary Enhanced edge Image*

896x592, is modified and reduced to a binary image with the size 324x194. The later one will undergo the process of Corner Detection.

ii. Corner Detection

The idea behind this approach is that each face, regading its oval shape, will form number of corners on lower resulation. To detect these corners on the left side of a face, we form a matrix of this form: $\begin{bmatrix} 0 & -1 & 0 \\ -1 & +1 & +1 \\ 0 & +1 & -1 \end{bmatrix}$ to detect all the regions of 9x9 pizels that include a L like shape constructed by three high bits surrounded by 6 low bits. This kernel is swiped on the whole original image from left to right, starting from upper-left of the edg-enhanced image. Once the described pattern is detected, the central pixel of the kernel is flagged as a starting point of the corner. This procedure leads to a binary image of the same size, with number of high pixels indicating detected left corner patter as shown in figure 3. This image then is investigated to detect the regions of highest density of high bits which will be described in the next subsection.

Figure 3: *Binary image, hoding the flags that indicate detected left corners*

iii. Detection of Interest Region

To decrease the size of our search space for a patch including a face, it is more efficient to focus on a region in which the probability of presence of a face is higher. For that purpose, we swip a vertical narrow rectangular with the same height as Edge image's. This rectangle is moved from left to right, calculating density of high bits inside its edges, then it flags regions with number of high bits above a treshold. This rectangular shapes are the regions we are interested in, and only the high bits trapped in these regions will be considered in following phase: Candidate Extraction.

iv. Candidate Extraction

So far, the precedure of focusing on a part of image has been discussed. In the Extraction level, we simply crop part of the original image with the help of information aquired by provious processings.Then it is fed to verifier, to see whether there is a face present in that pach or not. For Extraction purposes, we refer to the region of interest, map the pixels onto original image respecting resize scales and crop an image of fixed size with the high pixel locaing its upper-left corner. These candidate images, are feed to a machine learning based algorithm to be varified as a face.

v. Verification

The final step in this model, is to aknowledge the face that a candidate patch of the whole image, contains a face. For this purpose, we

have take advantage of a computationally efficient model based on Euclidean Distance in the terms of Machine learning. This component, is trained with a dataset, forming a matrix which holds each image of data set as a column. Candidate images is converted to a vector and it is checked weather it falls into a defined range.

IV. Results

Face detection is one of the most braod fields of study. Variant challenges lead to different approches which can be implemented with different perspectives and methods. We have developed a model with a focus on speed and low computational cost. Purposed model, takes an image, processes it in 3 phases to find a region of interest, then crops multiple patches and verfies the one with a face with the use of a distance based machin learning algorithm. We have used a single image, with multiple faces and a hormonous background. Wehave used Yale face database[34] to train machine learning component. Then the implementation is fed with 1. As the final result, method shows 79% accuracy (18 out of 23 faces in the pictures were correectly recognized.)

V. Evaluation and future works

This method is highly roboost against lightening, face expression and presence of facial features. However, it is dramatically sensitive to contrast of skin color with background and shape of faces appearing in the picture. Although results are not reflecting an extraordinary model, but the core idea is novel. The impementation has the potential of being improved, especially by extracting the best convolutional matrix for each enviroment using machine learning techniques.

References

[1] RAMA CHELLAPPA CHARLES L. AND SAAD SIROHEY **Human and Machine Recognition of Faces: A Survey**

[2] Ming-Hsuan Yang, David J. Kriegman, and Narendra Ahuja**Detecting Faces in Images: A Survey'**

[3] Mayank Chauhan et al, **'Study & Analysis of Different Face Detection Techniques.'**, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1615-1618

[4] G. Yang and T. S. Huang, **"Human Face Detection in Complex Background,"**, Pattern Recognition, vol. 27, no. 1, pp. 53-63, 1994.

[5] C. Kotropoulos and I. Pitas, **"Rule-Based Face Detection in Frontal Views"**, Proc. Int'l Conf. Acoustics, Speech and Signal Processing, vol. 4, pp. 2537-2540, 1997.

[6] K.C. Yow and R. Cipolla, **"A Probabilistic Framework for Perceptual Grouping of Features for Human Face Detection,"**, Proc. Second Int'l Conf. Automatic Face and Gesture Recognition, pp. 16-21, 1996.

[7] Y. Amit, D. Geman, and B. Jedynak, **"Efficient Focusing and Face Detection"**, Face Recognition: From Theory to Applications, H. Wechsler, P.J. Phillips, V. Bruce, F. Fogelman-Soulie, and T.S. Huang, eds., vol. 163, pp. 124-156, 1998.

[8] L. Breiman, J. Friedman, R. Olshen, and C. Stone, **"Classification and Regression Trees."**, Wadsworth, 1984.

[9] Y. Dai and Y. Nakano, **"Face-Texture Model Based on SGLD and Its Application in Face Detection in a Color Scene,"** Pattern Recognition, vol. 29, no. 6, pp. 1007-1017, 1996.

[10] R.M. Haralick, K. Shanmugam, and I. Dinstein, **"Texture Features for Image Classification,"** IEEE Trans. Systems, Man, and Cybernetics, vol. 3, no. 6, pp. 610-621, 1973.

[11] S. Satoh, Y. Nakamura, and T. Kanade, **"Name-It: Naming and Detecting Faces**

in News Videos," IEEE Multimedia, vol. 6, no. 1, pp. 22-35, 1999.

[12] J. Sobottka and I. Pitas, **"Segmentation and Tracking of Faces in Color Images,"** Proc. Second Int'l Conf. Automatic Face and Gesture Recognition, pp. 236-241, 1996.

[13] D. Chai and K.N. Ngan, **"Locating Facial Region of a Head-and- Shoulders Color Image,"** Proc. Third Int'l Conf. Automatic Face and Gesture Recognition, pp. 124-129, 1998.

[14] D. Saxe and R. Foulds, **"Toward Robust Skin Identification in Video Images,"** Proc. Second Int'l Conf. Automatic Face and Gesture Recognition, pp. 379-384, 1996.

[15] K. Sobottka and I. Pitas, **"Face Localization and Feature Extraction Based on Shape and Color Information,"** Proc. IEEE Int'l Conf. Image Processing, pp. 483-486, 1996.

.

[16] M.-H. Yang and N. Ahuja, **"Detecting Human Faces in Color Images,"** Proc. IEEE Int'l Conf. Image Processing, vol. 1, pp. 127-130, 1998

[17] T. Sakai, M. Nagao, and S. Fujibayashi, **"Line Extraction and Pattern Detection in a Photograph,"** Pattern Recognition, vol. 1, pp. 233-248, 1969.

[18] I. Craw, D. Tock, and A. Bennett, **"Finding Face Features,"** Proc. Second European Conf. Computer Vision, pp. 92-96, 1992.

[19] V. Govindaraju, **"Locating Human Faces in Photographs,"** Int'l J. Computer Vision, vol. 19, no. 2, pp. 129-146, 1996.

[20] V. Govindaraju, D.B. Sher, R.K. Srihari, and S.N. Srihari, **"Locating Human Faces in Newspaper Photographs,"** Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 549-554, 1989.

[21] V. Govindaraju, S.N. Srihari, and D.B. Sher, **"A Computational Model for Face Location,"** Proc. Third IEEE Int'l Conf. Computer Vision, pp. 718-721, 1990.

[22] T.F. Cootes and C.J. Taylor, **"Locating Faces Using Statistical Feature Detectors,"** Proc. Second Int'l Conf. Automatic Face and Gesture Recognition, pp. 204-209, 1996.

[23] M. Turk and A. Pentland, **"Eigenfaces for Recognition,"** J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[24] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss, **"The FERET Evaluation Methodology for Face-Recognition Algorithms,"** IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 10, pp. 1090-1034, Oct. 2000.

[25] F.S. Samaria, **"Face Recognition Using Hidden Markov Models,"** ,PhD thesis, Univ. of Cambridge, 1994.

[26] P. Hallinan, **"A Deformable Model for Face Recognition Under Arbitrary Lighting Conditions,"** PhD thesis, Harvard Univ., 1995.

[27] P. Belhumeur, J. Hespanha, and D. Kriegman, **"Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection,"** IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, 1997.

[28] A. Martinez and R. Benavente, **"The AR Face Database,"** Technical Report CVC 24, Purdue Univ., 1998.

[29] K.-K. Sung, **"Learning and Example Selection for Object and Pattern Detection,"** PhD thesis, Massachusetts Inst. of Technology, 1996.

[30] K.-K. Sung and T. Poggio, **"Example-Based Learning for View- Based Human Face Detection,"** Technical Report AI Memo 1521, Massachusetts Inst. of Technology AI Lab, 1994.

[31] H. Rowley, S. Baluja, and T. Kanade, **"Neural Network-Based Face Detection,"** Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 203-208, 1996.

[32] H.A. Rowley, **"Neural Network-Based Face Detection,"** PhD thesis, Carnegie Mellon Univ., 1999.

[33] A.C. Loui, C.N. Judice, and S. Liu, **"An Image Database for Benchmarking of Automatic Face Detection and Recognition Algorithms,"** Proc. IEEE Int'l Conf. Image Processing, pp. 146-150, 1998.

[34] Yale Face Database. urlhttp://vision.ucsd.edu/content/yale-face-database

[35] RAMA CHELLAPPA CHARLES L. AND SAAD SIROHEY **Human and Machine Recognition of Faces: A Survey**

[36] Ming-Hsuan Yang, David J. Kriegman, and Narendra Ahuja**Detecting Faces in Images: A Survey'**

[37] Mayank Chauhan et al, **'Study & Analysis of Different Face Detection Techniques.'**, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1615-1618

[38] Markus Weber, **"Frontal face dataset by California Institute of Technology"**, `http://www.vision.caltech.edu/archive.html`, access date: 2018-11-21