Acad. Year

2020-21

FACE RECOGNITION SYSTEM

Face Recognition System

Wardah Ali

Eesha Qureshi

Omama Ahmed Farooqi

Faculty of Information Technology

Department of Computer Science

Salim Habib University, Karachi, Pakistan

Year 2020/21

FACE RECOGNITION SYSTEM

Submitted by

Wardah Ali

(F18CSC05)

Eesha Qureshi

(F18CSC24)

Omama Ahmed Farooqi

(F18CSC26)

Faculty of Information Technology

Department of Computer Science

Machine Learning Course Project Report Presented to

Dr. Rizwan Ahmed Khan

Faculty of Information Technology

Department of Computer Science

Salim Habib University, Karachi, Pakistan

Year 2021

Contents

1. Introduction **1**
1.1. Face Detection 1
1.2. Feature Extraction 1
1.3. Face Recognition 2

2. Literature Review **3**
2.1. HAAR-Cascades(Viola Jones Method): 3
2.2. Histogram of Oriented Gradient 3
2.3. Principal Component Analysis 3
2.4. Gabor Filter 3
2.5. Support Vector Machine: 4

3. Methodologies **5**
3.1. HAAR-Cascade 5
3.2. Histogram Equalization 6
3.2..1 Experimental Results 6
3.3. Principal Component Analysis 7
3.4. 2-D Gabor Filter 7
3.5. Histogram of Oriented Gradient(HOG) 8
3.6. Support Vector Machine(SVM Kernels) 9

4. Appendices **10**

5. Results and Conclusion **11**
5.1. CLASSIFICATION REPORT 12

6. Future Work **14**

7. References **15**

1. Introduction

Recognizing faces is a natural and usually an effortless process for us humans but it's application has remained difficult in the area of computer vision. Face recognition stands steadfast as one of the most challenging and irrepressible areas in computer vision as it has evolved drastically and has been widely adopted in different bio-metric forms. It is broadly used in authentication tasks like identifying individuals, surveillance and access control based on features extracted from face images either statistically or geometrically. The crucial elements of a facial recognition system are as follows:

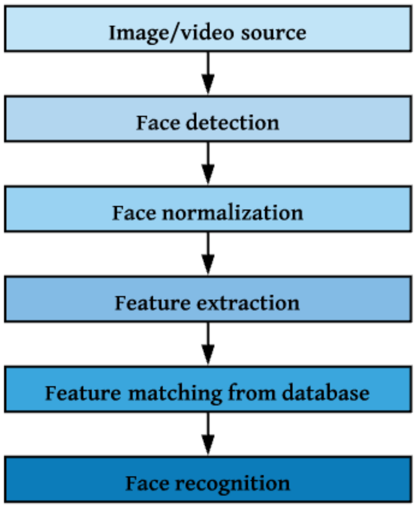


Figure 1.
elements of typical face recognition system

1.1. Face Detection

The first and most important function of a face recognition system is detection i.e whether faces appear in a given image/video or not and where these faces are located in them. Face detection uses algorithms and machine learning to find human faces in digital images/videos. This is done after some pre-processing that basically includes cropping the background to reduce the dimensionality. The output of face detection are patches that contain the human face of the input image. Face detection can be complicated due to a lot of factors such as lighting conditions, expressions, pose, image resolution etc.

1.2. Feature Extraction

After the face is detected and human-face patches are extracted from the input images, feature extraction is performed. Feature extraction can be done by a lot of algorithms. Feature extraction includes dimension reduction. Information packing, noise cleaning and salience extraction. A human-face patch is usually transformed to a vector with fixed dimensions.

1.3. Face Recognition

The last step is recognition of the human face. This requires a database that has several images of each person. Basically features are extracted from each class and stored in the database. When an input image comes in and after detection and extraction are performed, the features of the input image are matched with the features of the images present in the database. The process can be seen in the following figure:

2. Literature Review

2.1. HAAR-Cascades(Viola Jones Method):

Michael Jones and Paul Viola presented a framework for object detection which minimized computation time while achieving high detection accuracy. Michael Jones and Paul Viola**[1]** proposed a robust and fast method for detecting which was 15 times quicker than any techniques present at the time of release with 95 percent accuracy at around 17 fps. This technique relies on the use of HAAR-like features that are quickly evaluated through the usage of a new image representation. It is based on the concept of an integral image that generates a large set of features. It uses the boosting algorithm 'AdaBoost' to reduce the overcomplete set. The detector is applied on gray-scale images in a scanning-like fashion. The features evaluated can be scaled as well as the scanned window that is applied.

2.2. Histogram of Oriented Gradient

HOG or Histogram of Oriented Gradients is an image feature extractor that was first proposed by Dalal and Triggs **[2]** in 2005. The paper discusses feature extraction by HOG in-depth and how it can detect humans in images. As written in the paper, Histogram of Oriented Gradients is the preferred feature extractor for detecting human face. The paper discusses how HOG extracts global features instead of local features which makes it better at identifying faces or objects. Ebrahimzadeh and Jampour**[2]** in 2014 discussed how HOG is faster as compared to other descriptors such as LBP, SIFT etc. Furthermore, HOG is stable when it comes to illumination variation meaning it works efficiently in different lighting conditions.

2.3. Principal Component Analysis

Principal component analysis (PCA) is a way/method of data processing consisting of extractions of a number of synthetic variables, called principal components, from a huge number of variables calculated in order to explain a certain phenomenon. Principal components are projections of the data that are in sequential form. They are mutually ordered in variance and are uncorrelated. They are obtained as approximating linear manifolds a set of N points. PCA is considered a very useful tool for dimension compression and reduction as the resulted factors are orthogonal. Every factor of PCA explains a large part of the variation given by the N variables that satisfy a certain condition **[3]**.

2.4. Gabor Filter

The coefficients of Gabor are found to be outstanding for object detection and are robust to contrast differences and object distortions**[4]**. Gabor filter are basically matched to a line segment with a particular length that is applied to the image. After the filter output is integrated by connectivity of the trace region. The shorter Gabor wavelets are adapted to curved line traces and space variant intensity images. If the image has a space invariant the longer Gabor filter(kernel) loses such adaptability of the almost straight line traces and gray-level characteristic**[5]**. Gabor works robustly on morph/noisy images that's why we decided to use it in our system.

2.5. Support Vector Machine:

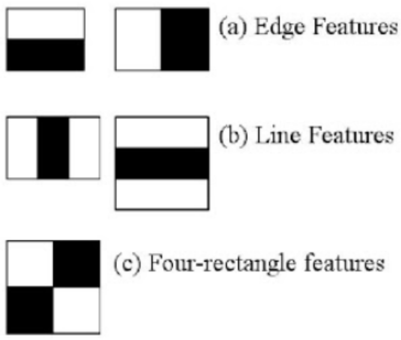
SVM was first introduced and used by Osuna et al. for face detection. Support Vector Machines work as a new paradigm to train linear, polynomial or radial basis function (RBF) classifiers. SVMs work on an induction principle, called structural risk minimization. Structural risk minimization targets to minimize upper bound on the expected generalization error. SVM is a linear classifier in which the separating hyper-plane is chosen to minimize the expected classification error of the unseen test data/patterns. Osuna et al.**[6]** invented an efficient method to train SVM for large scale problems and applied it to face detection. SVMs have also been used to detect pedestrians and faces in the wavelet domain.

3. Methodologies

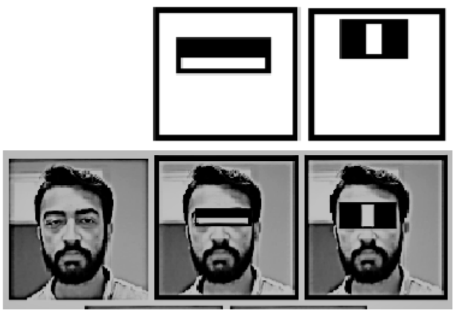
In this section, all the methodologies we tried using are defined and briefly explained. Due to hardware issues and keeping in mind the robustness of the system, we ended up using the following methodologies: (1) HAAR-Cascade (2) Histogram equalization (3) Gabor feature extraction (4) SVM classifier for classification

3.1. HAAR-Cascade

Detecting faces in images can be complicated due to the different variability factors such as expression, pose, position, orientation, pixel values, spectacles/glasses or facial hair, lighting conditions and resolution of the image. The Viola-Jones framework that was proposed in 2001 made major improvements in face detection methodology. The Viola Jones framework detects faces with high accuracy in real time. The basis of this is to train a model to recognise whether it's an actual face or not. Once it is trained, the specific features of the model are then extracted out. A file is created in which these new features are stored so that they can be compared with new images that were previously at various different stages.Viola-Jones framework uses HAAR-like features that makes find out the edges or lines in the image easy. HAAR-like features pick areas that have a sudden change in pixel intensity.
The three different types of HAAR features are shown in the figure below:



In the below picture it can be observed that the areas with higher pixel intensity that are the eyes have black boxes on them and the bridge of the nose which is the high point of the face with a lower pixel intensity has white box on it.



3.2. Histogram Equalization

Histogram equalization is a very important image enhancement method, which can improve the global contrast of the image. When the light in the environment of the face is too bright or too dark, the face image obtained by the camera will be blurred. However, under some circumstances histogram equalization might degrade the appearance of the image so we should always use care and judgment when applying this subjective method for image enhancement. In this system, the histogram equalization effectively improves the visual effect of the image before we do the feature extraction due to which we obtain the edge characteristics of our face image. The histogram of the image can be equalized using the equation:

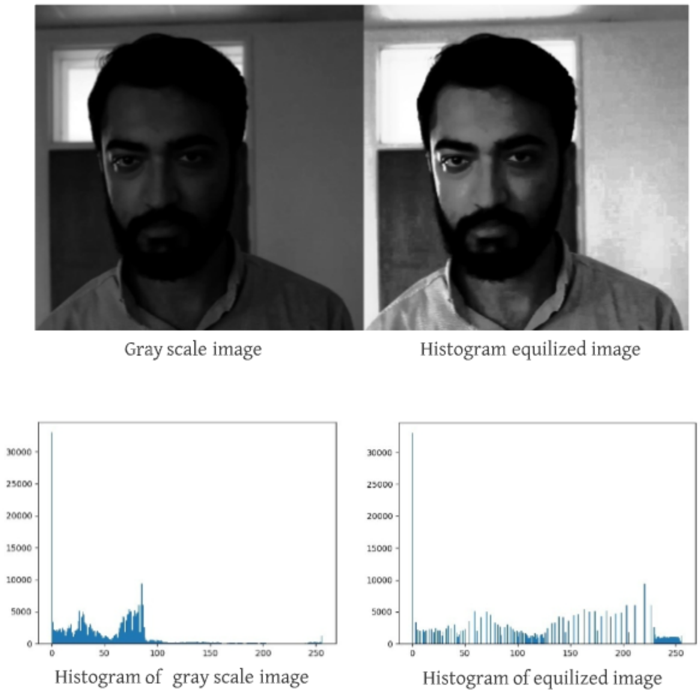
$$s_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j)$$

$$= \frac{L-1}{MN}\sum_{j=0}^{k} n_j\ ,\ k = 0, 1, 2, 3, \ldots., L-1$$

T (rk) is the transformation function that represents the grayscale mapping, rk is the kth gray scale value which is being mapped to a random variable sk. pr(rj) is the probability function of rj whose formula is written in which the number of pixels with intensity j (nj) are divided by the total number of pixels MN.

3.2..1 Experimental Results

Original Image RGB to gray and its Histogram Equalization result along with their histograms:



3.3. Principal Component Analysis

Gabor 2D Filter resulted in an inflation of the original pixel space, for image size of 256x256 we had the feature vector of 65537 which is too large. As The issue of storage space and computations are as important as the accuracy of verification and/or identification, so we implemented a technique to minimize the number of filters used by removing the redundancy and correlation between filters using Principal Component Analysis PCA.

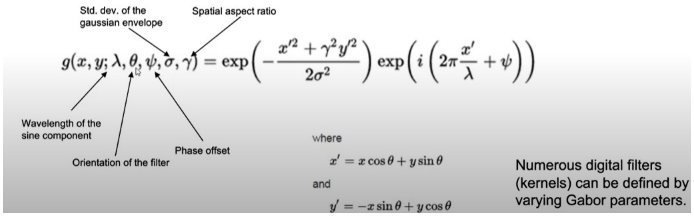
The PCA linearly transforms t-dimensional feature space into an f-dimensional feature subspace, where normally f«t. The new feature vector is defined by:

$$Y_i = W_{PCA}{}^{T} X_{i(i=1,2,\ldots N)}$$

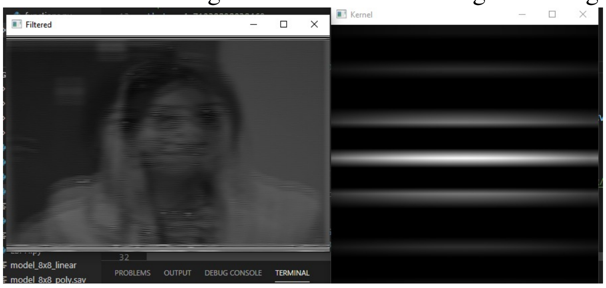
Principal components are constructed in such a manner that the first principal component accounts for the largest possible variance in the data set that is the component with the largest eigen value. The second principal component is calculated in the same way, with the condition that it is uncorrelated with (i.e., perpendicular to) the first principal component and that it accounts for the next highest variance.

3.4. 2-D Gabor Filter

Gabor filter is a linear filter used for edge detection In the spatial domain, a 2D Gabor filter is based on components of Gaussian function modulated by a sinusoidal plane wave. The general expression of the Gabor filter family is implemented as:



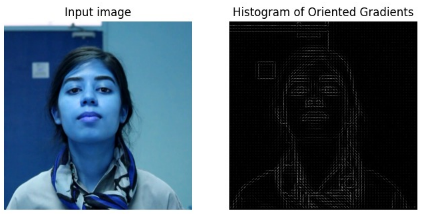
The image below on the left is a hidden image from our data set along with the gabor kernel applied on it.



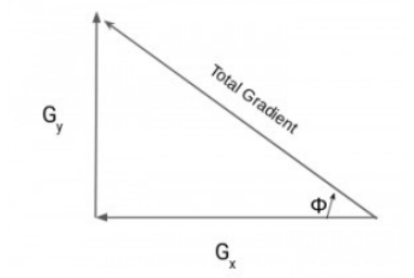
3.5. Histogram of Oriented Gradient(HOG)

Initially, before using the Gabor feature extraction method we used HOG. The Histogram of Oriented Gradients method is mainly utilized for face and image detection to classify images. It is widely used in computer vision tasks for object detection. We had 2691 images originally of 1080x1080 PX. After detecting the faces the images were cropped and we resized it to 256x256 PX.

HOG acts as a feature descriptor that simply extracts the useful information and disregards the unnecessary information from the image. These images were perfect for image classification algorithms, like in our case we used SVMs in order to produce good results. Now, in order for the HOG feature descriptor to sort



through the unnecessary information it calculates the gradient for every pixel in the image. Gradients are extremely important for checking for corners and edges in an image especially in the areas where there are many intensity changes. The gradients are calculated in both x and y directions separately, this is done for all the pixels in the image. The next step would be to find the magnitude and orientation using these values. The magnitude and the direction of the gradients can simply be found by using the Pythagoras theorem: To move on to the next step of the HOG algorithm, make sure that the image is divided into cells so that the histogram of gradients can be calculated for each cell. We divided our images into 8x8 cells by doing so, we got the features (histogram) for the smaller patches which in turn represent the whole image. In the end we combine all the features to get to the final image.



Magnitude: $g = \sqrt{g_x^2 + g_y^2}$

Direction: $\theta = \arctan \frac{g_y}{g_x}$

THE HISTOGRAMS ARE CALCULATED BY USING THE GRADIENTS AND ORIENTATION

3.6. Support Vector Machine(SVM Kernels)

LINEAR KERNEL: A linear kernel can be used as a normal dot product of any two given observations. The product between two vectors is the sum of the multiplication of each pair of input values. The linear kernel works fine when the data set is linearly separable. In our case since we had a large number of features in the data set, linear kernel was used. Training a SVM with a Linear Kernel is faster than with any other Kernel. POLYNOMIAL KERNEL: A polynomial kernel is a more generalized form of the linear kernel.

$$K = y*x + b$$

In the polynomial kernel, we simply calculate the dot product by increasing the power of the kernel. Here

$$K(X_1, X_2) = (a + X_1^T X_2)^b$$

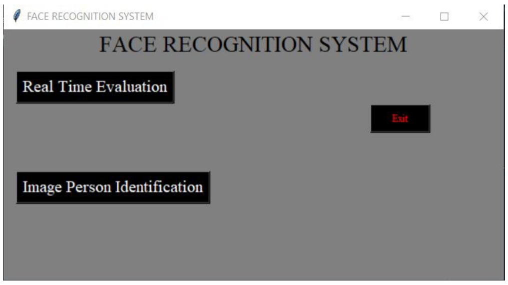
b is the degree of polynomial and 'a' is the constant term. RBF KERNEL: Gaussian RBF (Radial Basis Function) is a function whose value depends on the distance from the origin or from some point. RBF can map an input space in infinite dimensional space. Gamma () controls the influence of new features, a higher

$$K(X_1, X_2) = exponent(-\gamma\|X_1 - X_2\|^2)$$

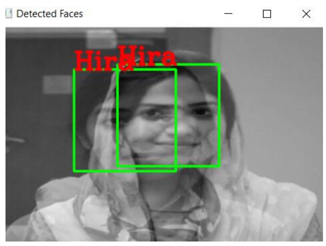
value of gamma will perfectly fit the training dataset (Gamma=0.1 is considered to be a good default value) K(X1, X2) on the decision boundary. The higher the gamma, the more influence the features will have on the decision boundary. ||X1 - X2||2 is the Euclidean distance between X1 and X2.

4. Appendices

The Graphical User Interface of our project looks like:



The main reason for using Gabor filter is that it works remarkable on morph images as you can see in the picture below:



The following is an image of our face recognition system in working.

5. Results and Conclusion

We applied HOG on the cropped images of the faces, accuracy results were good but it didn't work at all on noisy data so we had to drop it. The alternate method we used to extract features is 2D Gabor Filter

Gabor and HOG kernel size=5	Train Accuracy	Test Accuracy
SVM Linear kernel	90%	80.3%
SVM Polynomial kernel	50%	42.3%
SVM RBF kernel	45.6%	39.8%

with kernel size 9 (as it preferred to use for normal images) and trained it on different SVM kernels. The accuracy results it gave are as follows:

Gabor kernel size=9	Train Accuracy	Test Accuracy
SVM Linear kernel	99.8%	42.8%
SVM Polynomial kernel	59.17%	30.01%
SVM RBF kernel	58.18%	30%

Since the accuracy results were not up-to the mark at kernel size 9, we changed the kernel to size 5 and then trained the model again, This time the results were as follows:

Gabor kernel size=5	Train Accuracy	Test Accuracy
SVM Linear kernel	100%	98.3%
SVM Polynomial kernel	99.4%	98.5%
SVM RBF kernel	99.5%	98.14%

After applying a 2D Gabor filter on the dataset, it still had the shape of 2691x65537. It was too large to be trained and we were facing curse of dimensionality, so we applied PCA.

Gabor and PCA kernel size=5	Train Accuracy	Test Accuracy
SVM Linear kernel	100%	98.3%
SVM Polynomial kernel	99.8%	98.1%
SVM RBF kernel	99.6%	98.8%99

Gabor along with PCA gave really good accuracies but when we ran the model file on run time, it wasn't efficient and we were facing hardware issues so we had to drop PCA. Therefore, we only applied histogram equalization and used 2D Gabor Filter with the SVM linear kernel which gave us fairly good results. Although our program is not efficient enough to recognize faces from noisy data we surely do look forward to working on it.

5.1. CLASSIFICATION REPORT

In the following picture precision is showing the percentage of true positives, which is is 100 percent in most classes. Recall measures the completeness of the classifier. It finds the number of true positives by the ratio of true positives to the sum of true positives and false negatives, mathematically it is written as: TP/(TP+FN) F1-Score is a weighted harmonic mean of precision and recall such that the best score is 1.0 and the worst is 0.0. Generally speaking, F1 scores are lower than accuracy measures as they embed precision and recall into their computation. Support is the number of actual occurrences of the class in the specified data set.Imbalanced supports shows structural weakness that we look forward to work on.

Classification Report

	precision	recall	f1-score	support
Abdullah	1.00	0.96	0.98	24
Affan	0.96	1.00	0.98	25
Ali	1.00	0.97	0.98	32
Aziz	0.96	1.00	0.98	27
Basit	1.00	1.00	1.00	14
Eesha	1.00	1.00	1.00	23
Eman	1.00	0.92	0.96	26
Faraz	1.00	1.00	1.00	30
Fasih	1.00	0.97	0.99	34
Hala	1.00	1.00	1.00	36
Hamra	0.88	1.00	0.93	14
Hasan	1.00	1.00	1.00	27
Hira	0.97	0.97	0.97	36
Jamali	0.85	1.00	0.92	17
Jawwad	0.93	1.00	0.96	27
Laviza	1.00	0.96	0.98	25
Parshant	1.00	0.97	0.98	31
Rehmat	1.00	1.00	1.00	29
Shehriyar	1.00	1.00	1.00	13
Subhan	1.00	1.00	1.00	24
Wardah	1.00	0.92	0.96	25
accuracy			0.98	539
macro avg	0.98	0.98	0.98	539
weighted avg	0.98	0.98	0.98	539

Here's the confusuion matrix for linear SVM showing that the majority of the dataset falls in the true positives or true negatives.

```
Confusion Matrix

[[22  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 24  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0 30  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0 34  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0 26  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 15  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  1  0  0  0  0 22  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0 31  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0 33  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0 25  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 11  0  0  1  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0 21  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 39  0  0  0  0  0  0  0  0]
 [ 0  1  0  0  0  0  0  1  0  0  0  0  0 32  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0 30  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 18  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 25  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 27  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 24  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 21  0]
 [ 0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0 23]]
```

6. Future Work

We will work on increasing the robustness and accuracy of our system on noisy data. Moreover, we'll try to incorporate this system in our coming semester projects/FYP. If we choose to do research, face recognition system will open doors for a lot of new ideas.

7. References

[1] P. Viola, M. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features, 2001.

[2]N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection.INRIA Rhone-Alps, ^ 655 avenue de l'Europe, Montbonnot 38334, France

[3] Hastie, T., Tibshirani, R., Friedman, J.: The elements of statistical learning. Data mining, Inference, and prediction. 2nd edition. Springer, 2009.

[4] Shuzo Yamamoto et al. "Extraction of fluorescent dot traces from a scanning Laser Opthalmoscope Image sequence by Spatio-Temporal Image analysis: Gabor Filter and Radon Transform Filtering" pp.1357-1363, vol. 46, IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, 1999

[5] David P. Casasent, John Scott Smokelin, Anqi Ye, "Optical Gabor, wavelet, and morphological filters for image processing," Proc. SPIE 1772, Optical Information Processing Systems and Architectures IV, (12 January 1993).

[6] Osuna et al. (2000). Training support vector machines: an application to face detection. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Published.

https://www.datacamp.com/

https://www.medium.com/