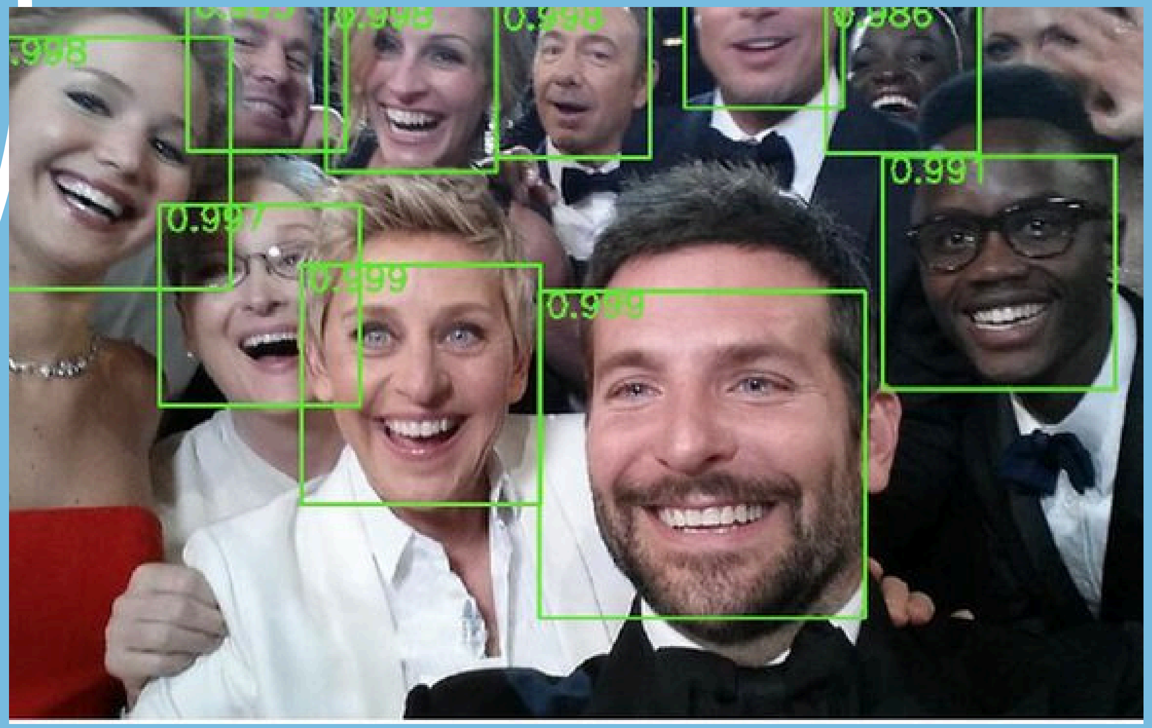
Ministry of Science and Higher Education of the Republic of Kazakhstan
L.N. Gumilyov Eurasian National University

Faculty of Information Technology
Department of Information Systems

VIOLA-JONES METHOD

Done by: Zhumabek A.R

INTRODUCTION

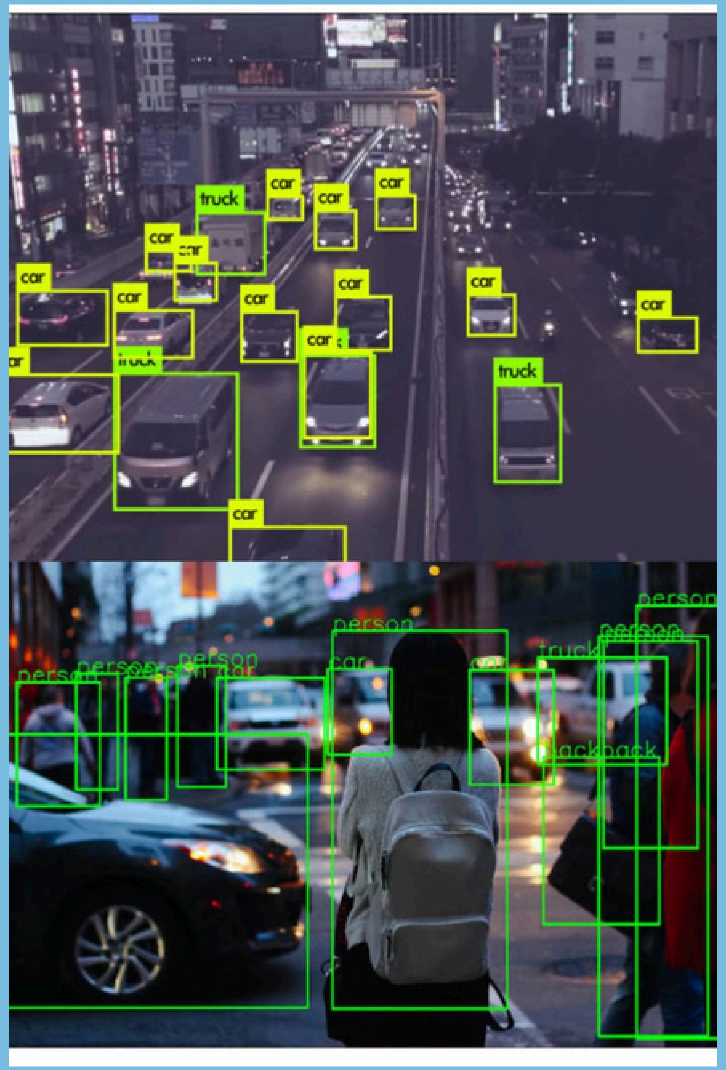


Face recognition is one of the most studied tasks in such fields as digital image processing, computer vision, biometrics, videoconferencing, intelligent security and access control systems, and so on. The process of face recognition usually consists of two steps: searching for the face region in the image and comparing the found face with the faces in the database. Currently, the Viola-Jones method is the most popular method for searching for the face region in an image due to its high speed and efficiency.

The method was developed and presented in 2001 by Paul Viola and Michael Jones. It is still an effective method for finding objects in real-time images and video sequences [1, 2]. It should be noted that this detector has an extremely low false face detection probability. The method performs well and detects facial features even when the object is observed at a small angle, up to about 30°.

THE VIOLA-JONES IS BASED ON SEVERAL KEY PRINCIPLES

01	Haar-like features
02	Integral Image
03	Adaboost training
04	Cascade Classifiers



HAAR-LIKE FEATURES

In terms of the need for simple feature acquisition algorithms, the use of Haar like features, which are the result of comparing the brightnesses in two rectangular regions of an image, is promising. It turns out Haar-like features are simple patterns used to represent contrasting regions of an image. They help in identifying key structures such as eyes, nose, mouth and facial contours.

The standard Viola-Jones method uses rectangular features, Figure 1. These features are called Haar primitives.

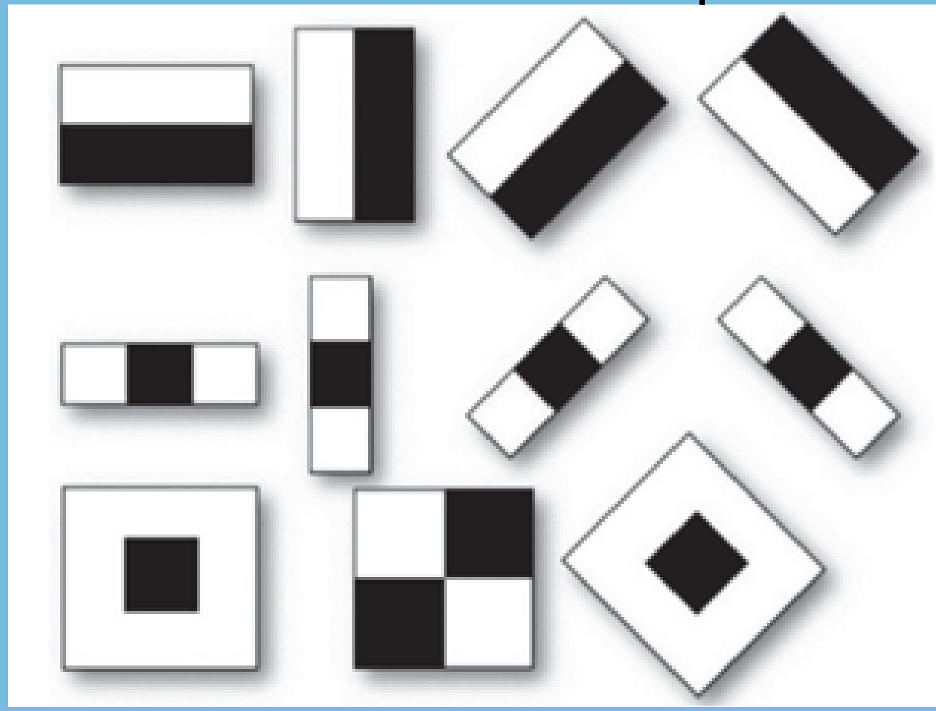


Fig. 1. Haar feature primitives

The extended Viola-Jones method, available in the OpenCV library, uses additional features (Fig. 2)

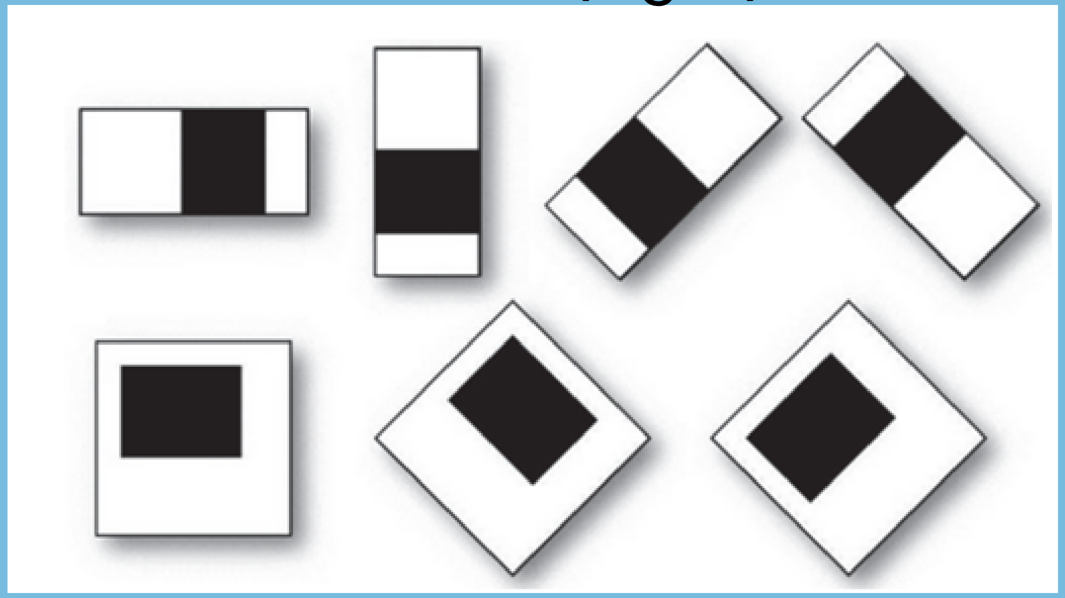


Fig.2. Additional Haar features

The computed value of such a feature will be: (Where U is the sum of the brightness values of the points covered by the light part of the feature, and V is the sum of the brightness values of the points covered by the dark part of the feature. The concept of integral image is used for their calculation.)

INTEGRAL IMAGE

In order to calculate the brightness of a rectangular section of an image, an integral representation is used [3]. This representation is often used in other methods such as wavelet transforms, Speeded up robust feature(SURF), Haar filters and many developed algorithms. Integral Image allows to quickly calculate the sum of pixels in any rectangular section of the image.

Integral image representation is a matrix that matches the size of the original image. Each element stores the sum of intensities of all pixels to the left and above the element. Matrix elements are calculated by the following formula:

$$I(x,y) = \sum_{x' \leq x,\, y' \leq y} i(x',y'),$$

where I(x,y) is the value of the point (x,y) of the integral image; i(x,y) is the intensity value of the original image.

The integral representation has an interesting feature. By the integral matrix it is possible to calculate very quickly the sum of pixels of an arbitrary rectangle.

ADABOOST TRAINING

Boosting is a technique that improves the accuracy of machine learning models by building a consistent composition of algorithms. The idea behind boosting is to combine weak classifiers, each of which individually produces results that differ only slightly from random guessing, into one strong classifier. This approach was proposed by Robert Schapire (Schapire) in the late 1990s.

Viola-Jones method uses AdaBoost (Adaptive Boosting) to strengthen the classifier and select the most significant Haar features among the many possible features. The process of building a classifier goes through several steps:

01

Weak classifiers are created. Each weak classifier is based on one Haar-attribute and determines whether there is a face in the domain under consideration.

02

The most significant features are selected. The algorithm analyzes which classifiers give the least number of errors and increases their weights.

03

A strong classifier is formed. The final classifier is a combination of several weak classifiers, where each of them is assigned a certain weight depending on its accuracy.

04

Building a cascade of classifiers. A face is passed through a cascade of filters, where at each step obviously incorrect candidates are discarded. This reduces the number of areas that require detailed analysis, which significantly speeds up the face detection process.

CASCADE CLASSIFIERS

The cascade structure increases the detection rate by focusing on the most informative areas of the image. The structure of the cascade detector is shown in Fig. 3.

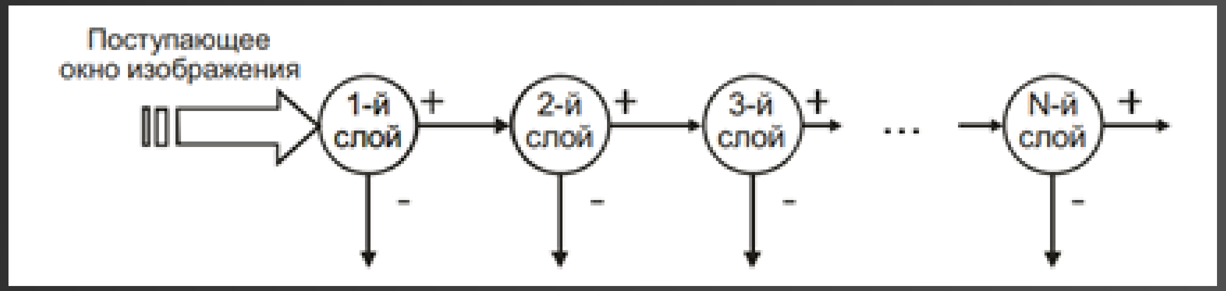


Fig. 3. Structure of the cascade detector

The cascade consists of successive layers, where each layer performs filtering: early layers quickly discard unwanted regions, while later layers perform more accurate analysis. If a region fails to pass at least one layer, it is immediately excluded, reducing computational cost.

Each layer is a strong classifier trained using bousting (AdaBoost). Early layers use fewer features for fast performance, while later layers analyze only promising regions. During the process, the first layer performs coarse filtering, the second layer analyzes the remaining regions in more detail, and the final layers make the final decision about the presence of a face.

The key advantages of the cascade classifier are high speed due to early rejection of inappropriate regions, saving computational resources and gradual increase in processing complexity. With this structure, the Viola-Jones method can run in real time, remaining one of the fastest face detection algorithms.

WAVELET TRANSFORMS

Wavelet Transform is a mathematical tool for analyzing signals and images, allowing to represent data at different scales. It is widely used in image processing, data compression, noise reduction and other tasks where it is important to analyze both local and global features of a signal.

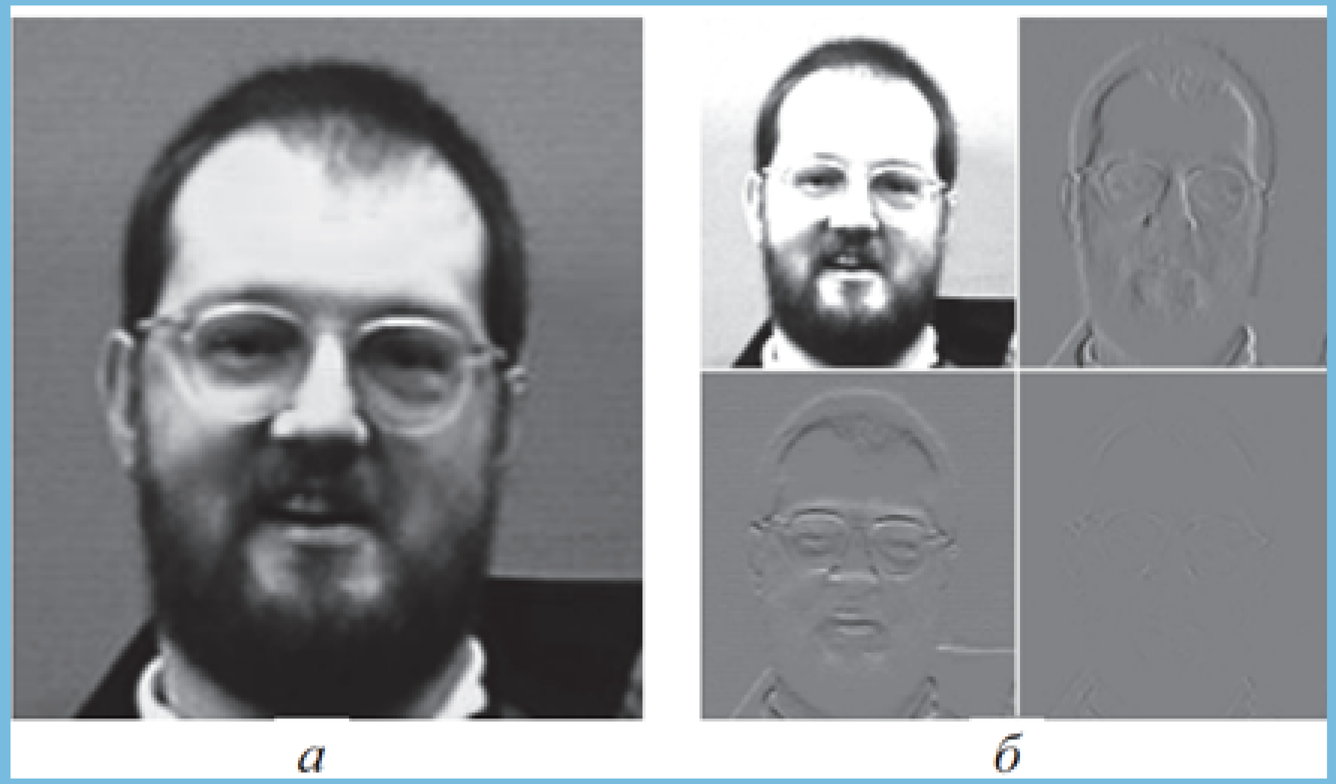


Figure 4. Example of face feature extraction: a) original face image; b) result after applying the Haar wavelet transform

- Popular wavelets

✔ Haar wavelet - the simplest, used for edge detection and sharp changes.

✔ Dobeshi wavelet - has a good compromise between signal localization and frequency analysis.

✔ Morlet wavelet - similar to a sinusoid with a damped envelope, used in speech and audio signal processing.

Examples of application of Haar wavelet transform and Dobeshi wavelet transform to extract features of a face image are shown in Fig. 4.

The code loads a detector for cat faces that knows what they look like. It then opens the image.Using the Viola-Jones method, the muzzles are searched for, given their size and position. Green rectangles are drawn around the found muzzles. The image with frames is then shown on the screen.

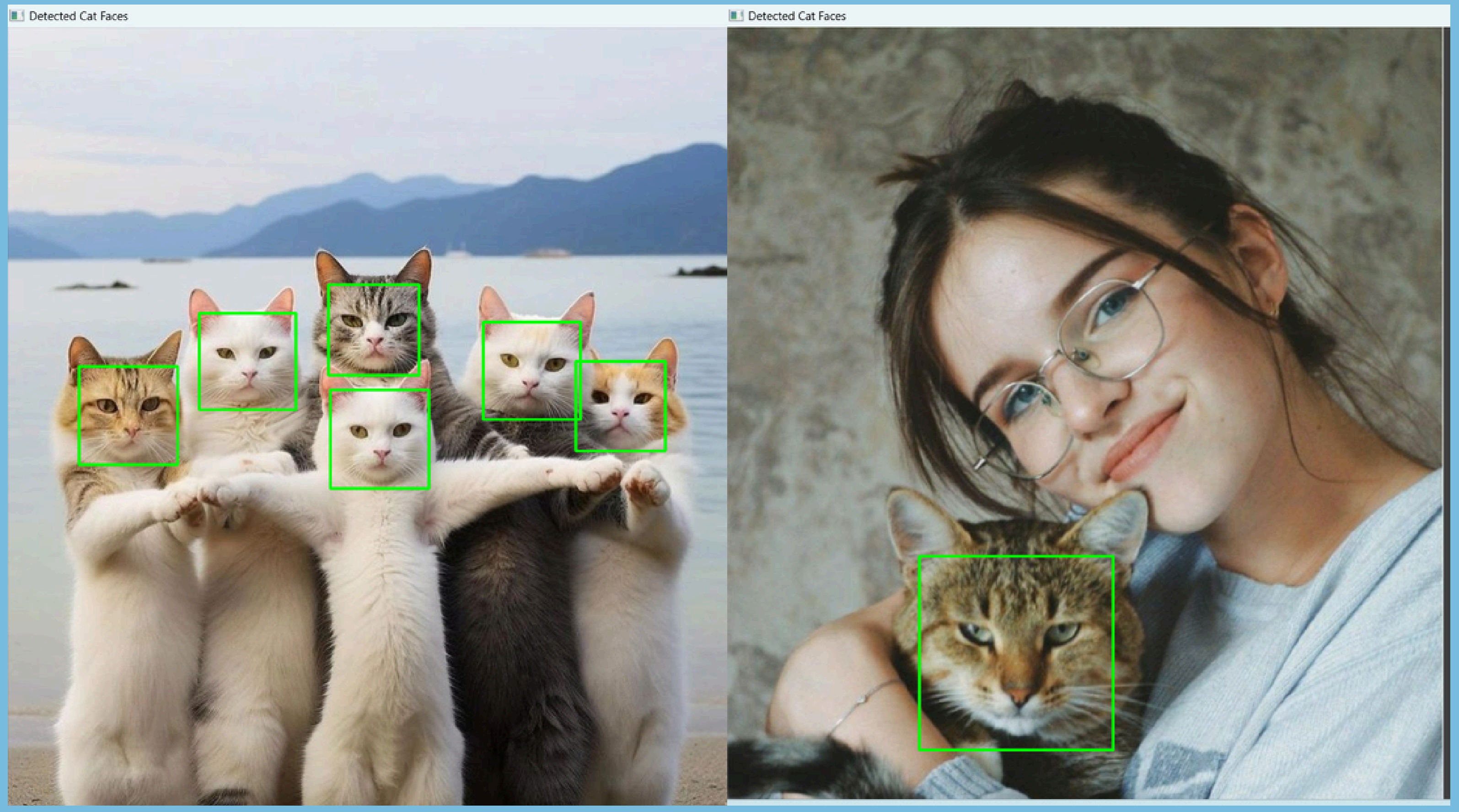
```
import cv2

cat_cascade = cv2.CascadeClassifier(cv2.data.haarcascades + 'haarcascade_frontalcatface.xml')
image = cv2.imread('C:/Users/User/Desktop/0c887018d152ecadae0f637c06de1048.jpg')
if image is None:
    print("mistake with import")
else:
    gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
    cats = cat_cascade.detectMultiScale(gray, scaleFactor=1.1, minNeighbors=5)
    for (x, y, w, h) in cats:
        cv2.rectangle(image, (x, y), (x+w, y+h), (0, 255, 0), 2)

    cv2.imshow( winname: 'Detected Cat Faces', image)
    cv2.waitKey(0)
    cv2.destroyAllWindows()
```

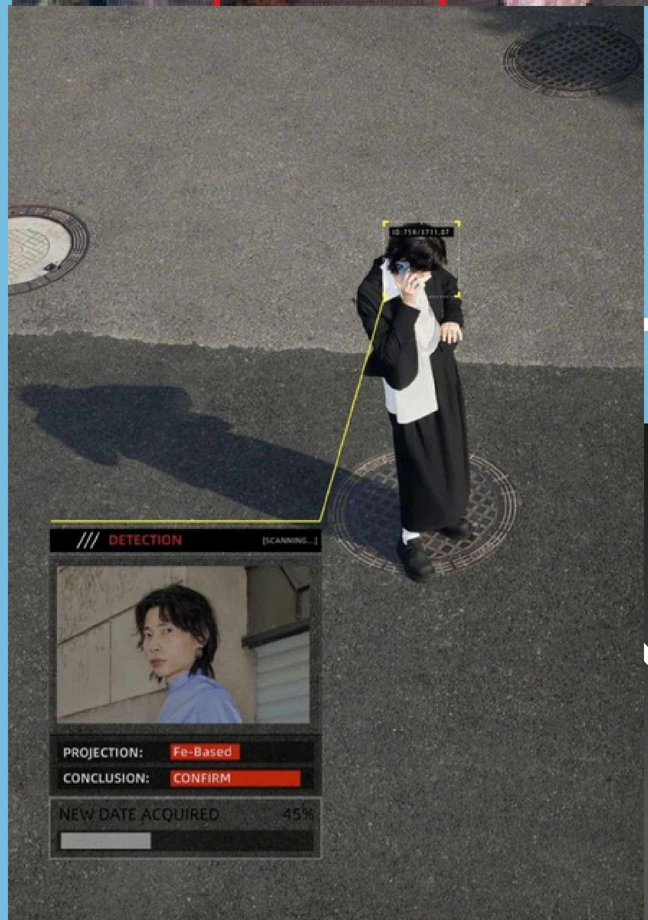
This code uses the Viola-Jones method to find cat faces in an image.
How it works:

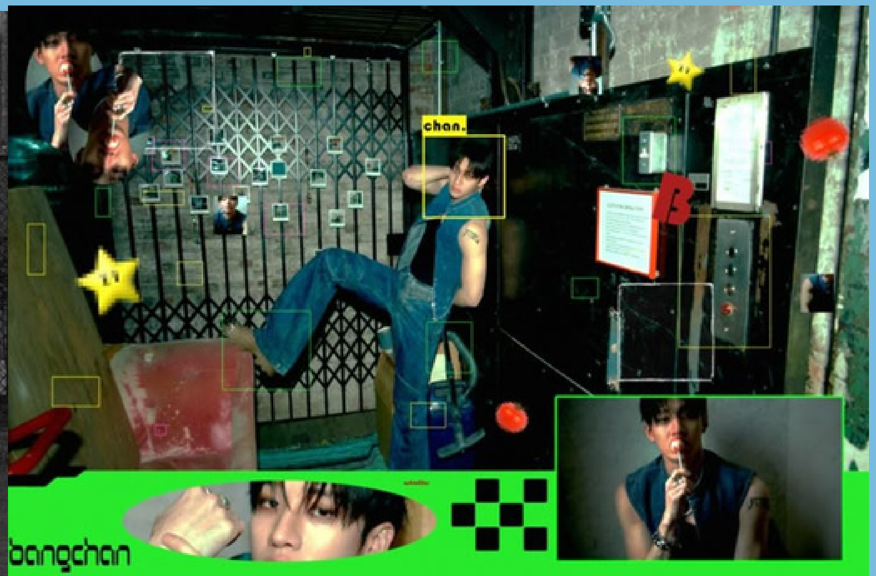
1. An image is loaded and converted to grayscale.
2. The Viola-Jones cascade searches for cat faces on it.
3. It finds areas where there are muzzles and draws green rectangles around them.
4. The finished image is shown on the screen.



CONCLUSION

Final scheme of work of the Viola-Jones method
🔷 1. Integral image - speeds up calculations.
🔷 2. Haar signs - analyzes key facial structures.
🔷 3. AdaBoost - selects the most relevant features.
🔷 4. Classifier Cascade - discards false positives and speeds up detection.
📌 Through these steps, the Viola-Jones method is able to find faces in images in real time and remains one of the fastest face detection algorithms. 🚀





THANK YOU

References

1. www.reallygreatsite.com
2. https://www.ipol.im/pub/art/2014/104/article.pdf
3. https://medium.com/0xcode/the-viola-jones-face-detection-algorithm-3eb09055cfc2