

Face Detection Using Machine Learning

Submitted By: Satyam Kumar

Course: CSE1021

Date: 20/11/2025

Introduction:

Face detection is a computer vision technique used to identify and locate human faces in images or videos. It serves as a foundation for applications like facial recognition, surveillance, emotion detection, and human–computer interaction. Machine learning algorithms enable efficient and accurate detection by learning patterns from large datasets.

Problem Statement:

Traditional manual monitoring systems struggle to detect and analyze faces accurately. The goal is to develop an automated face detection system using machine learning that can reliably identify human faces in real-time or static images.

Functional Requirements:

- The system must detect human faces in input images.
- The system should draw bounding boxes around detected faces.
- It should process real-time video streams (optional).
- It must support multiple face detection in one frame.

Non-functional Requirements:

- Accuracy: High detection accuracy under different lighting conditions.
- Performance: Real-time or near real-time processing.
- Scalability: Should handle multiple images or video frames.
- Usability: Easy to integrate into applications.

System Architecture:

- Input image/video
- Preprocessing module
- Machine learning-based face detection model (e.g., Haar Cascade, CNN-based detectors)
- Output module to visualize detection results

Code:

```
1 import imutils
2 import numpy as np
3 import cv2
4 from google.colab.patches import cv2_imshow
5 from IPython.display import display, Javascript
6 from google.colab.output import eval_js
7 from base64 import b64decode
```

```
1 def take_photo(filename='photo.jpg', quality=0.8):
2     js = Javascript('''
3         async function takePhoto(quality) {
4             const div = document.createElement('div');
5             const capture = document.createElement('button');
6             capture.textContent = 'Capture';
7             div.appendChild(capture);
8
9             const video = document.createElement('video');
10            video.style.display = 'block';
11            const stream = await navigator.mediaDevices.getUserMedia({video: true});
12
13            document.body.appendChild(div);
14            div.appendChild(video);
15            video.srcObject = stream;
16            await video.play();
17        }
18    
```

```
18     google.colab.output.setIframeHeight(document.documentElement.scrollHeight, true);
19
20     await new Promise((resolve) => capture.onclick = resolve);
21
22     const canvas = document.createElement('canvas');
23     canvas.width = video.videoWidth;
24     canvas.height = video.videoHeight;
25     canvas.getContext('2d').drawImage(video, 0, 0);
26     stream.getVideoTracks()[0].stop();
27     div.remove();
28     return canvas.toDataURL('image/jpeg', quality);
29 }
30 ''')
31 display(js)
32 data = eval_js('takePhoto().format(quality)')
33 binary = b64decode(data.split(',')[1])
34 with open(filename, 'wb') as f:
35     f.write(binary)
36 return filename
```

```
1 image_file = take_photo()

1 image = cv2.imread(image_file)
2 image = imutils.resize(image, width=400)
3 (h, w) = image.shape[:2]
```

```
5 cv2_imshow(image)

1 !wget -N https://raw.githubusercontent.com/opencv/opencv/master/samples/dnn/face_detector/deploy.prototxt
2 !wget -N https://raw.githubusercontent.com/opencv/opencv_3rdparty/dnn_samples_face_detector_20170830/res10_300x300_ssd_iter_140000.caffemodel
```

```
1 print("[INFO] loading model...")
2 prototxt = 'deploy.prototxt'
3 model = 'res10_300x300_ssd_iter_140000.caffemodel'
4 net = cv2.dnn.readNetFromCaffe(prototxt, model)

1 image = imutils.resize(image, width=400)
2 blob = cv2.dnn.blobFromImage(cv2.resize(image, (300, 300)), 1.0, (300, 300), (104.0, 177.0, 123.0))

1 print("[INFO] computing object detections...")
2 net.setInput(blob)
3 detections = net.forward()

1 for i in range(0, detections.shape[2]):
2     confidence = detections[0, 0, i, 2]
3     if confidence > 0.5:
4         box = detections[0, 0, i, 3:7] * np.array([w, h, w, h])
5         (startX, startY, endX, endY) = box.astype("int")
6         text = "{:.2f}%".format(confidence * 100)
7         v = startY - 10 if startY - 10 > 10 else startY + 10
```

A screenshot of a code editor window. The code is written in Python and uses the OpenCV library. It defines a function that takes an image and a bounding box as input. Inside the function, it converts the bounding box to integers, calculates a text position, draws a rectangle on the image, and adds text to the image. Finally, it displays the image using cv2.imshow. The code is as follows:

```
5  (startX, startY, endX, endY) = box.astype("int")
6  text = "{:.2f}%".format(confidence * 100)
7  y = startY - 10 if startY - 10 > 10 else startY + 10
8  cv2.rectangle(image, (startX, startY), (endX, endY), (0, 0, 255), 2)
9  cv2.putText(image, text, (startX, y),
10   cv2.FONT_HERSHEY_SIMPLEX, 0.45, (0, 0, 255), 2)
```

The status bar at the bottom shows the command "1 cv2_imshow(image)".

Workflow:

1. Load image
2. Preprocess
3. Apply ML detection model
4. Output results

Sequence Diagram:

User → System → ML Model → System → Output

Component Diagram:

- Image Processor
- Face Detector
- Display Module

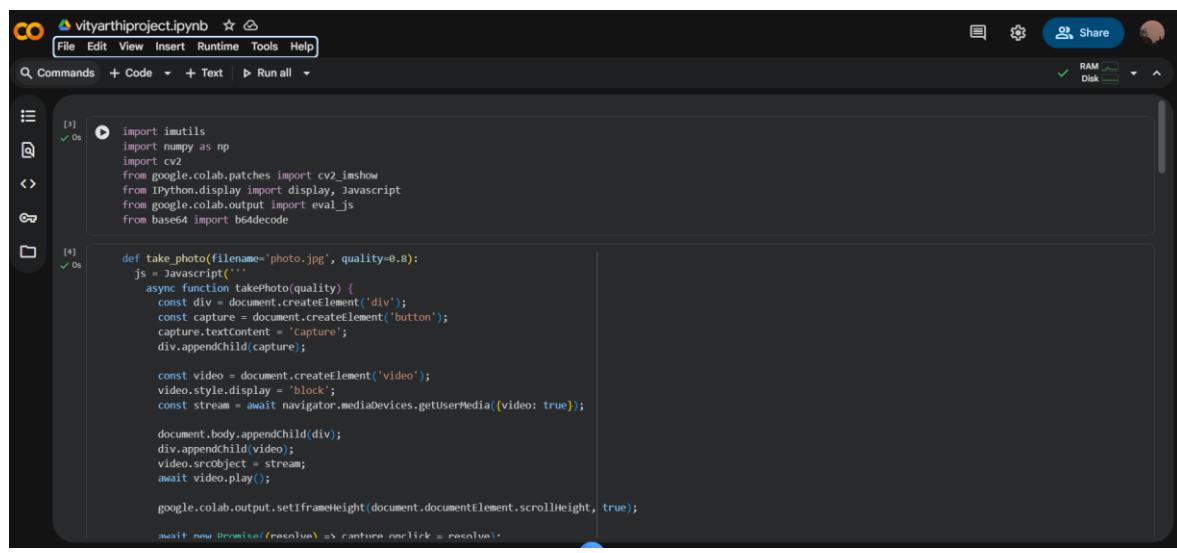
Design Decisions & Rationale:

- Python and OpenCV used due to wide library support.
- Modular architecture ensures reusability.

Implementation Details:

- Programming Language: Python
- Libraries: OpenCV, NumPy

Results:



The screenshot shows a Jupyter Notebook cell with two parts. Part [3] contains Python imports and a function definition. Part [4] contains a block of JavaScript code. The Python code includes imports for imutils, numpy, cv2, and various Google Colab patches. It defines a function `take_photo` that takes a filename and quality as arguments. The JavaScript part creates a div, a button, and a video element, sets up a media stream, and plays the video. It also handles a click event to capture the photo.

```
[3]
import imutils
import numpy as np
import cv2
from google.colab.patches import cv2_imshow
from IPython.display import display, Javascript
from google.colab.output import eval_js
from base64 import b64decode

[4]
def take_photo(filename='photo.jpg', quality=0.8):
    js = Javascript('''
        async function takePhoto(quality) {
            const div = document.createElement('div');
            const capture = document.createElement('button');
            capture.textContent = 'Capture';
            div.appendChild(capture);

            const video = document.createElement('video');
            video.style.display = 'block';
            const stream = await navigator.mediaDevices.getUserMedia({video: true});

            document.body.appendChild(div);
            div.appendChild(video);
            video.srcObject = stream;
            await video.play();

            google.colab.output.setFrameHeight(document.documentElement.scrollHeight, true);

            await new Promise((resolve) => capture.onclick = resolve);
        }
    ''')
    display(js)
    eval_js('takePhoto({})'.format(quality))
```

vityarthiproject.ipynb

```
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all

[4] ✓ 0s
    google.colab.output.setFrameHeight(document.documentElement.scrollHeight, true);

    await new Promise((resolve) => capture.onclick = resolve);

    const canvas = document.createElement('canvas');
    canvas.width = video.videoWidth;
    canvas.height = video.videoHeight;
    canvas.getContext('2d').drawImage(video, 0, 0);
    stream.getVideoTracks()[0].stop();
    div.remove();
    return canvas.toDataURL('image/jpeg', quality);

}
display(js)
data = eval_js(`takePhoto(${i})`.format(quality))
binary = b64decode(data.split(',')[1])
with open(filename, 'wb') as f:
    f.write(binary)
return filename

[9] ✓ 34s
image_file = take_photo()
```

vityarthiproject.ipynb

```
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all

[10] ✓ 0s
    image = cv2.imread(image_file)
    image = imutils.resize(image, width=400)
    (h, w) = image.shape[:2]
    print(w,h)
    cv2.imshow(image)

400 300

```

vityarthiproject.ipynb

```
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all

[11] ✓ 1s
    !wget -N https://raw.githubusercontent.com/opencv/opencv/master/samples/dnn/face_detector/deploy.prototxt
    !wget -N https://raw.githubusercontent.com/opencv/opencv_3rdparty/dnn_samples_face_detector_20170830/res10_300x300_ssd_iter_140000.caffemodel
    - 2025-11-23 16:40:36 - https://raw.githubusercontent.com/opencv/opencv/master/samples/dnn/face_detector/deploy.prototxt...
    Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 185.199.108.133, 185.199.109.133, 185.199.110.133, ...

[12] ✓ 0s
    image = imutils.resize(image, width=400)
    blob = cv2.dnn.blobFromImage(cv2.resize(image, (300, 300)), 1.0, (300, 300), (104.0, 177.0, 123.0))

[13] ✓ 0s
    print("[INFO] computing object detections...")
    net.setInput(blob)
    detections = net.forward()
    ...
    [INFO] computing object detections...

[14] ✓ 0s
    for i in range(0, detections.shape[2]):
        confidence = detections[0, 0, i, 2]
        if confidence > 0.5:
            box = detections[0, 0, i, 3:7] * np.array([w, h, w, h])
            (startX, startY, endX, endY) = box.astype("int")
            text = "{}: {:.2f}%".format(confidence * 100, confidence * 100)
            y = startY - 10 if startY - 10 > 10 else startY + 10
            cv2.rectangle(image, (startX, startY), (endX, endY), (0, 0, 255), 2)
            cv2.putText(image, text, (startX, y), cv2.FONT_HERSHEY_SIMPLEX, 0.45, (0, 0, 255), 2)
```

vityarthiproject.ipynb

```
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all

[15] ✓ 1s
    cv2.imshow(image)
    ...

```

Testing Approach:

- Test with multiple images under varying lighting conditions.
- Evaluate detection accuracy and false positives.
- Performance testing for real-time video.

Challenges Faced:

- Low light affecting detection accuracy.
- Handling occluded faces.
- Real-time processing speed.

Learnings & Key Takeaways:

- Understanding computer vision concepts.
- Hands-on experience with OpenCV.
- Improved understanding of ML model performance.

Future Enhancements:

- Use deep learning models (e.g., MTCNN, SSD).
- Add facial recognition.
- Deploy as a web or mobile app.

References::

- OpenCV Documentation
- Machine Learning textbooks
- Research papers on face detection
- <https://www.geeksforgeeks.org/machine-learning/face-recognition-using-artificial-intelligence/>
- <https://amanxai.com/2020/11/17/face-mask-detection-with-machine-learning/>
- <https://pyimagesearch.com/2020/05/04/covid-19-face-mask-detector-with-opencv-keras-tensorflow-and-deep-learning/>
- https://en.wikipedia.org/wiki/Face_detection