Face Detection

Aric Sharma 210906166
Akshaj Jauhri 210907446

11:11PM

Agenda

Topics Covered

- Introduction
- Objective
- Dataset
- Methodology

Start

nda · Agenda · Agenda · Agenda · Agenda · Agenda · Agenda · Agenda · Agenda · Agenda · Agenda · Agenda · Agenda · Agen

Agenda

Topics Covered

Start

Results

Findings

Conclusion

References

INTRODUCTION

Back to Agenda Page

Exploring Face Detection Algorithm Performance

Explore four distinct face detection models: Haar Cascade, DNN Caffe Model, HoG Model, and DNN YOLO Model.

Delve into the unique advantages and limitations of each model, encompassing computational efficiency and accuracy under varying conditions.

Conduct a comparative analysis to unveil performance differences, examining factors such as accuracy, speed, and robustness, to equip the user with insights for informed algorithm selection.

Back to Agenda Page

Haar Cascade

Classic, lightweight algorithm detects objects, including faces, using predefined features and cascading classifiers.

Pros

Simple and fast with low computational requirements

Cons

Limited accuracy in challenging conditions, like occlusions and varying lighting. Ineffective for non-frontal images, prone to false positives.

DNN Caffe

Deep neural network model for face detection trained using the Caffe framework, offering high accuracy but requiring significant computational resources.

Pros

High accuracy, especially with large and diverse datasets, can handle challenging conditions better than traditional methods.

Cons

Higher computational requirements compared to Haar Cascade and requires a trained model and a more complex setup.

HoG Model

Detects object boundaries, including faces, with moderate accuracy and lower computational requirements compared to deep learning models.

Pros

Moderate accuracy, lower computational needs than DNN models. Effective for frontal and slightly non-frontal faces. Lightweight compared to other models.

Cons

Doesn't detect small faces, trained for min size of 80x80. Less accurate than deep learning models, for complex backgrounds. Slower than Haar Cascade.

DNN YOLO

Employs YOLO architecture for efficient face detection by simultaneously predicting bounding boxes and class probabilities, with speed and accuracy.

Pros

High accuracy and robustness. Faster inference compared to other deep learning models due to its single-pass architecture.

Cons

Higher computational requirements compared to traditional methods.
May require more training data and fine-tuning for optimal performance.

OBJECTIVE

Back to Agenda Page

The objective of this project is to meticulously assess the performance of diverse face detection algorithms through comprehensive comparison.

Through meticulous evaluation, we aim to compare the accuracy, speed, and robustness of each algorithm under various conditions and scenarios.

By analyzing these factors, we intend to determine the suitability of each algorithm for different applications, ranging from real-time systems to high-resolution image processing.

Ultimately, our goal is to provide valuable insights that empower users to make informed decisions when selecting the most appropriate face detection algorithm for their specific needs and constraints.

Comprehensive Evaluation of Face Detection

Back to Agenda Page

DATASET

Back to Agenda Page

Crowd Dataset

Back to Agenda Page

Expressions Datatset

Back to Agenda Page

Faces Dataset

Back to Agenda Page

Occlusion Dataset

Back to Agenda Page

METHODOLOGY

Back to Agenda Page

odology • Methodology • Methodology • Methodology • Methodology • Methodology • Methodology • Methodology • Methodolo

During the comparison of models using a small dataset, we assess the following criteria

Accuracy

Compare the accuracy of each model based on benchmarks and real-world performance.

Speed

Discuss the inference speed of each model, considering real-time applications.

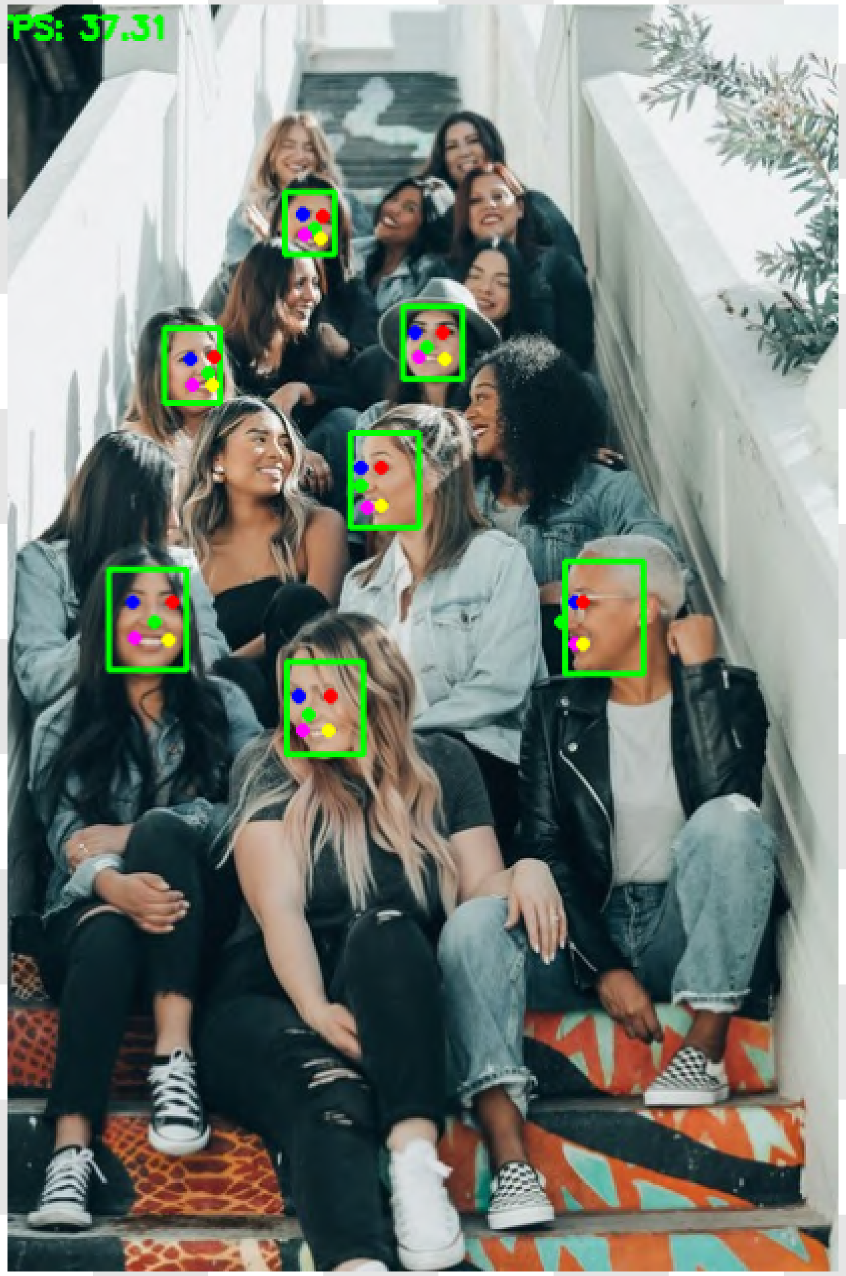
Robustness

Evaluate how each model performs under conditions like varying lighting, occlusions, and scale variations.

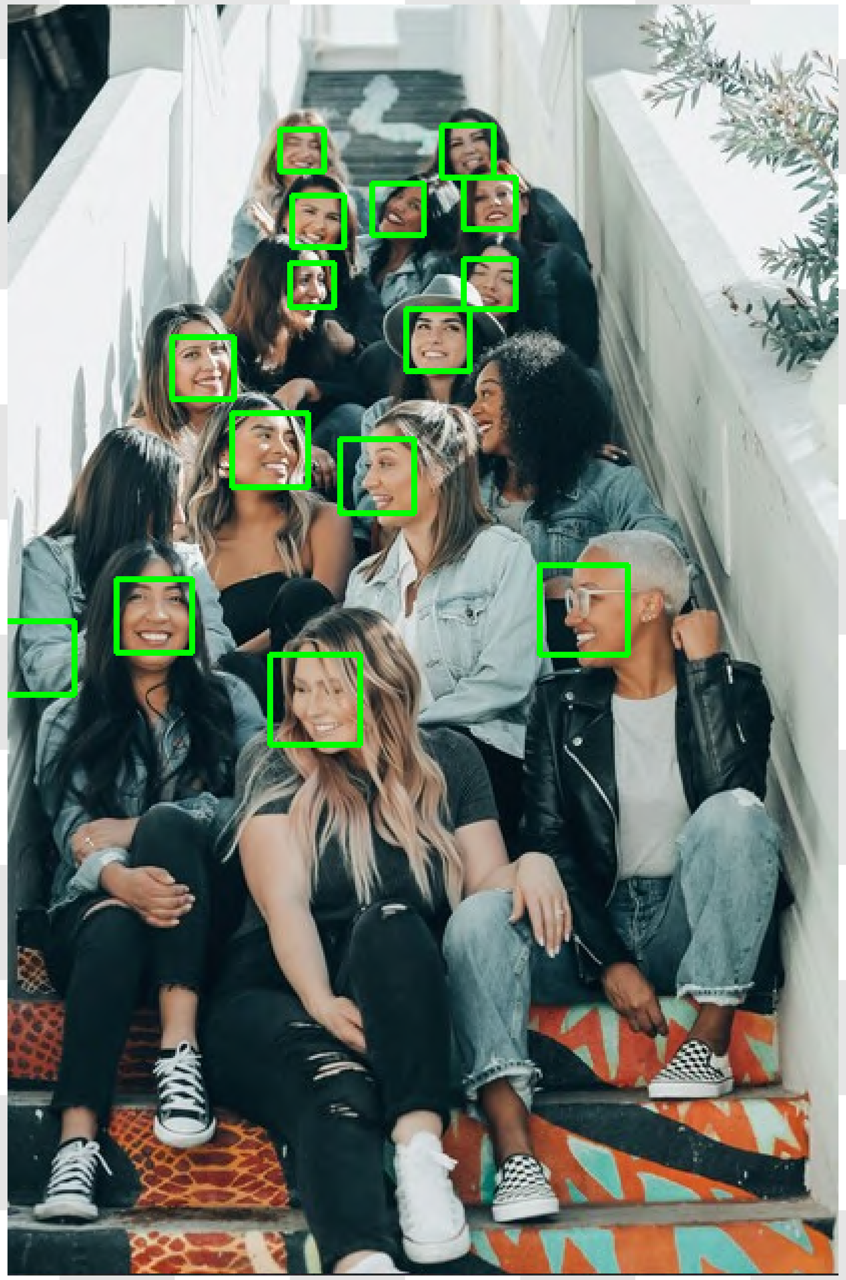
RESULTS

Back to Agenda Page

Haar Cascade



DNN Caffe



HoG

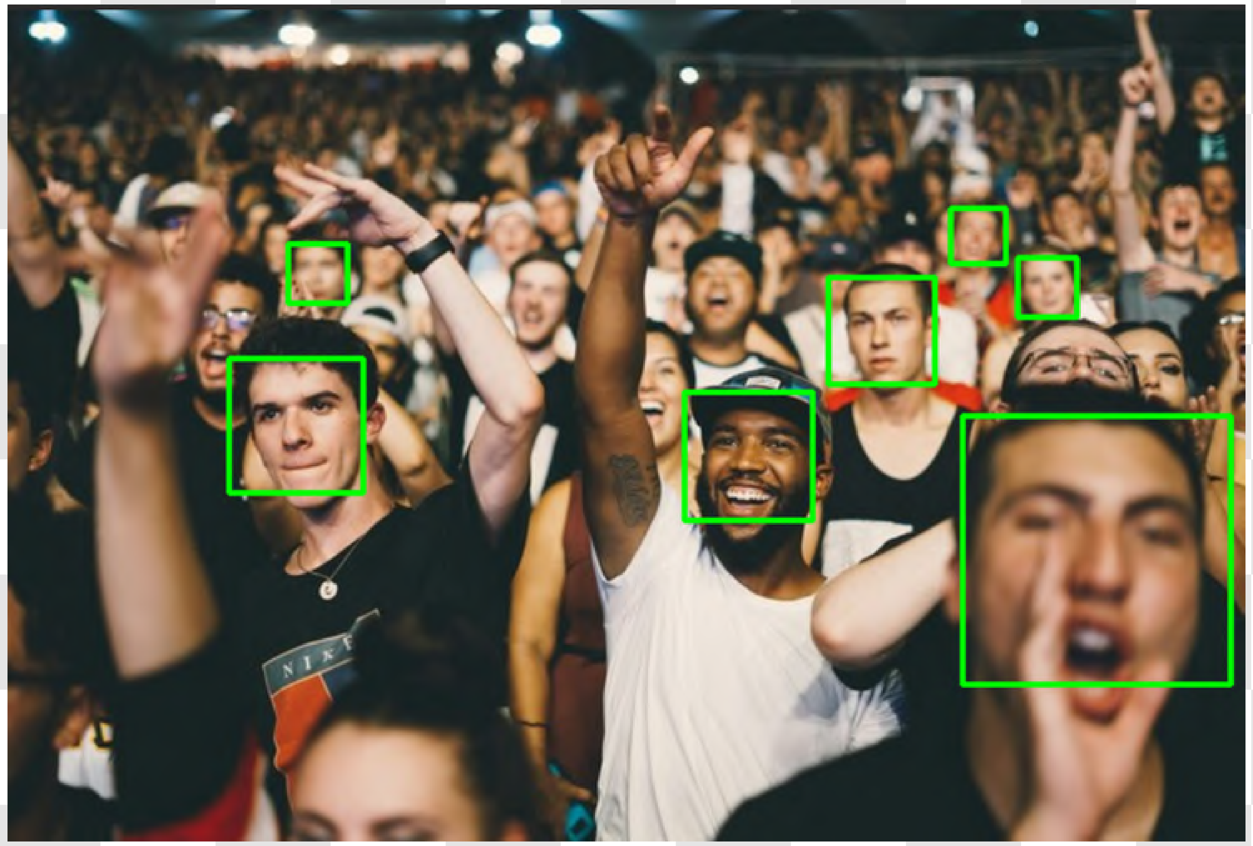
DNN YOLO

Haar Cascade

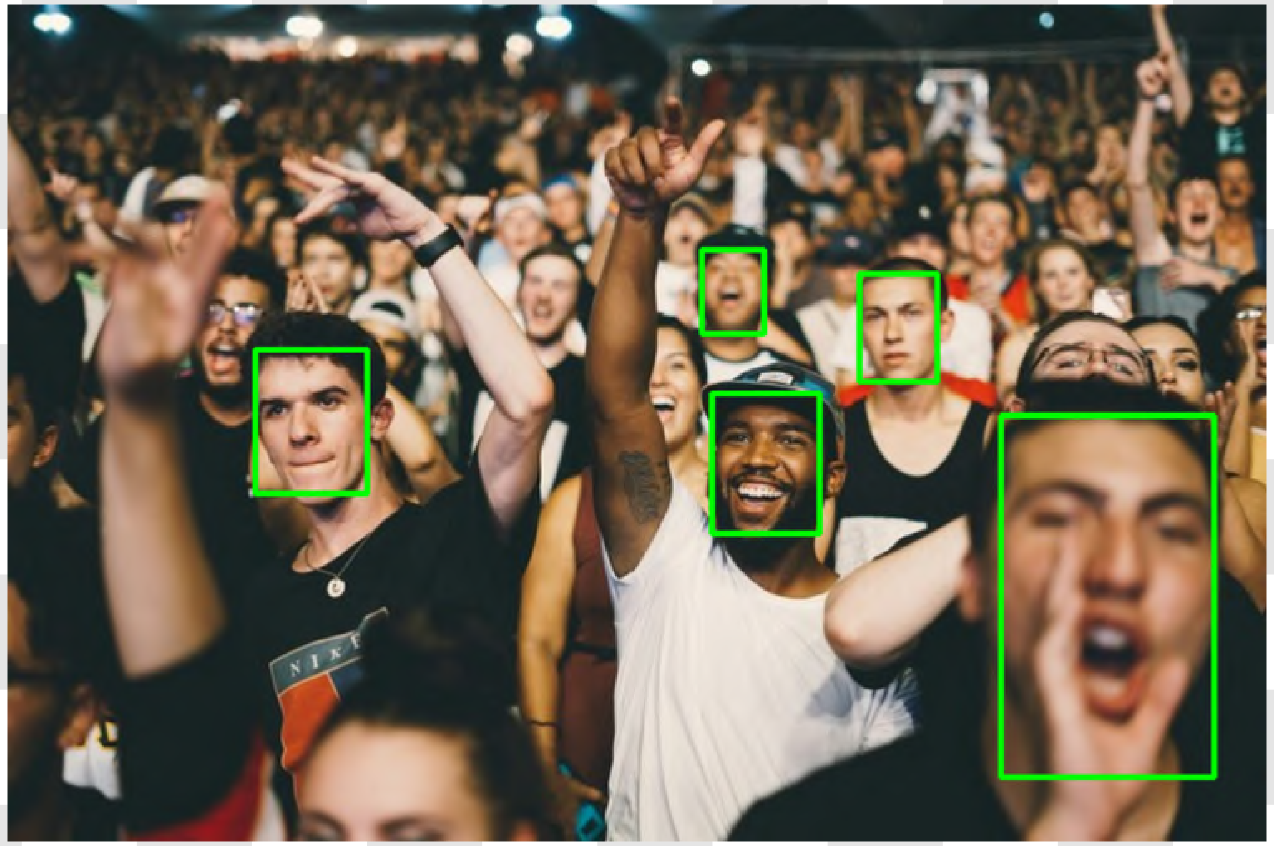
DNN Caffe

HoG

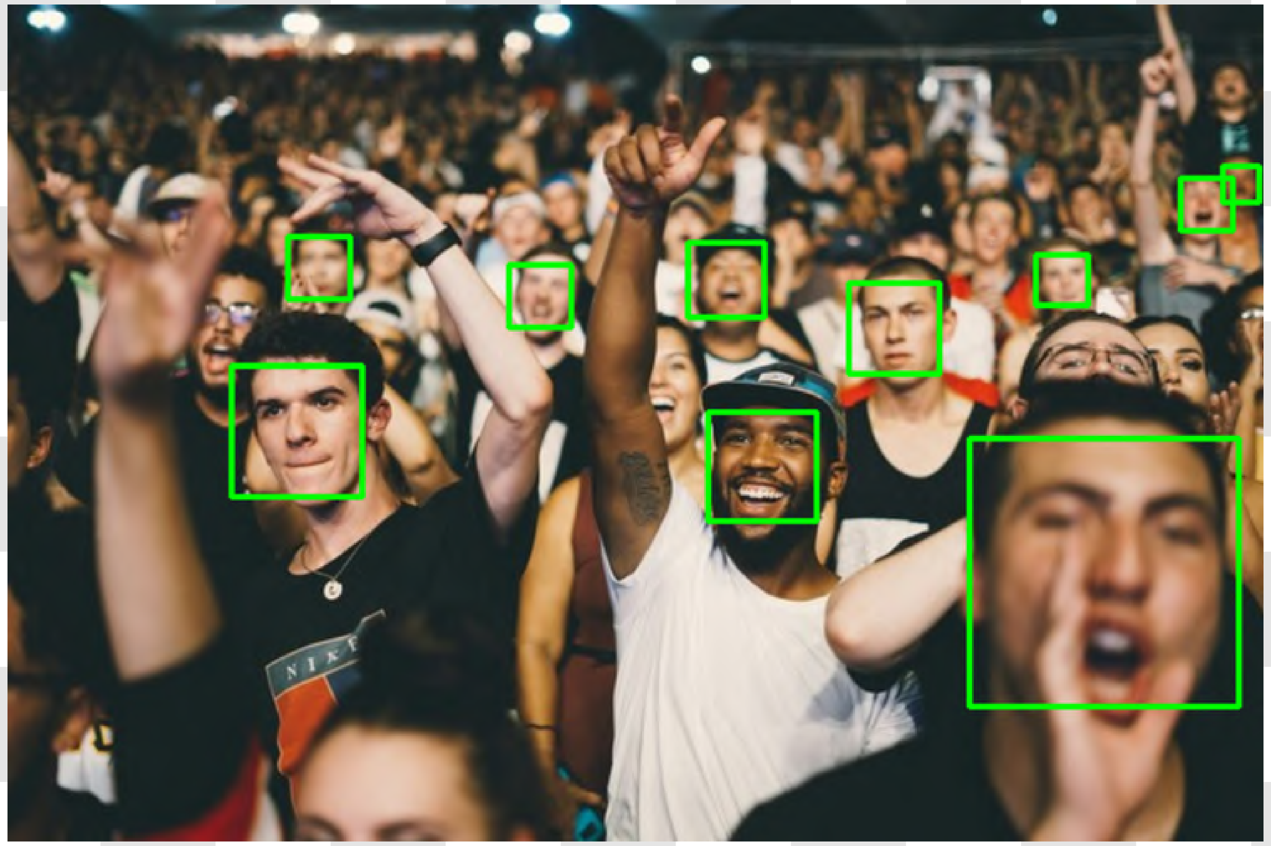
DNN YOLO



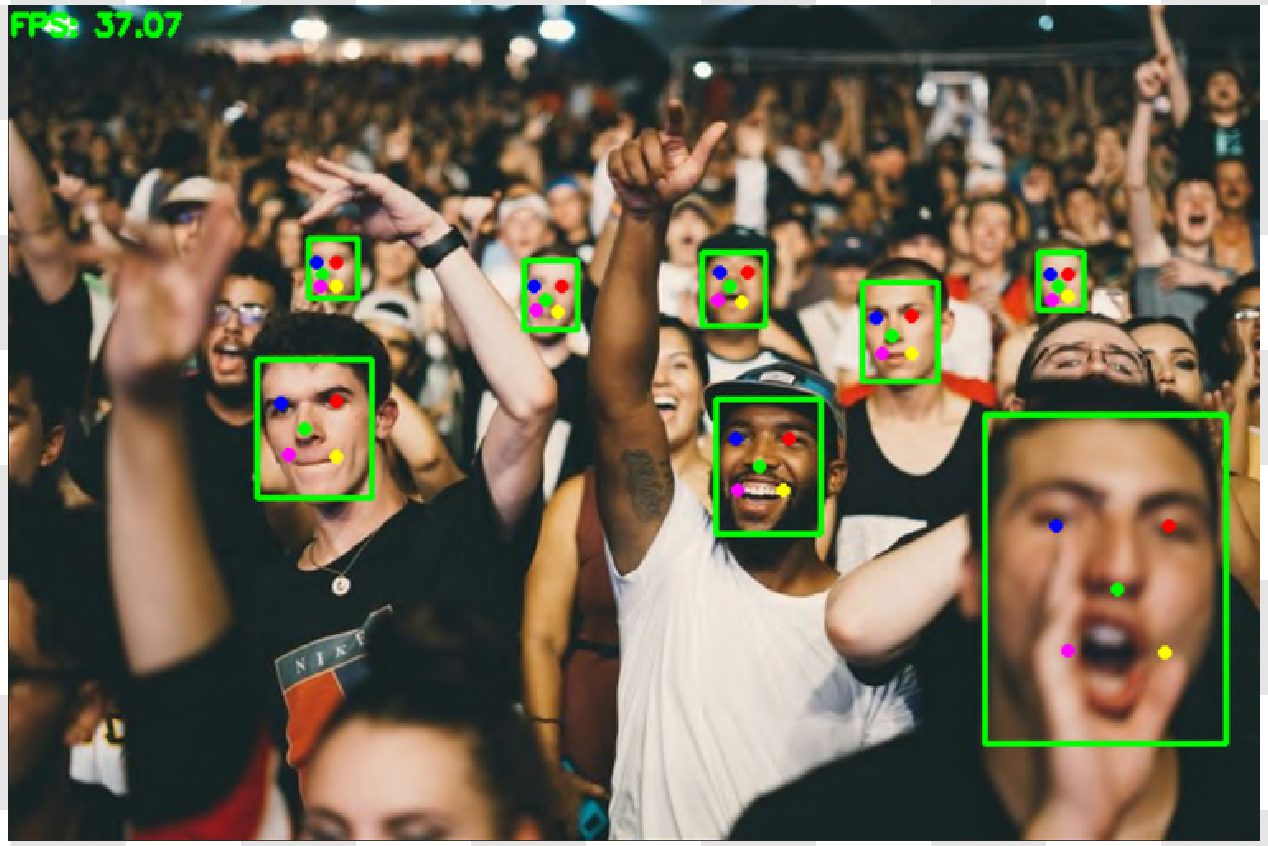
Haar Cascade



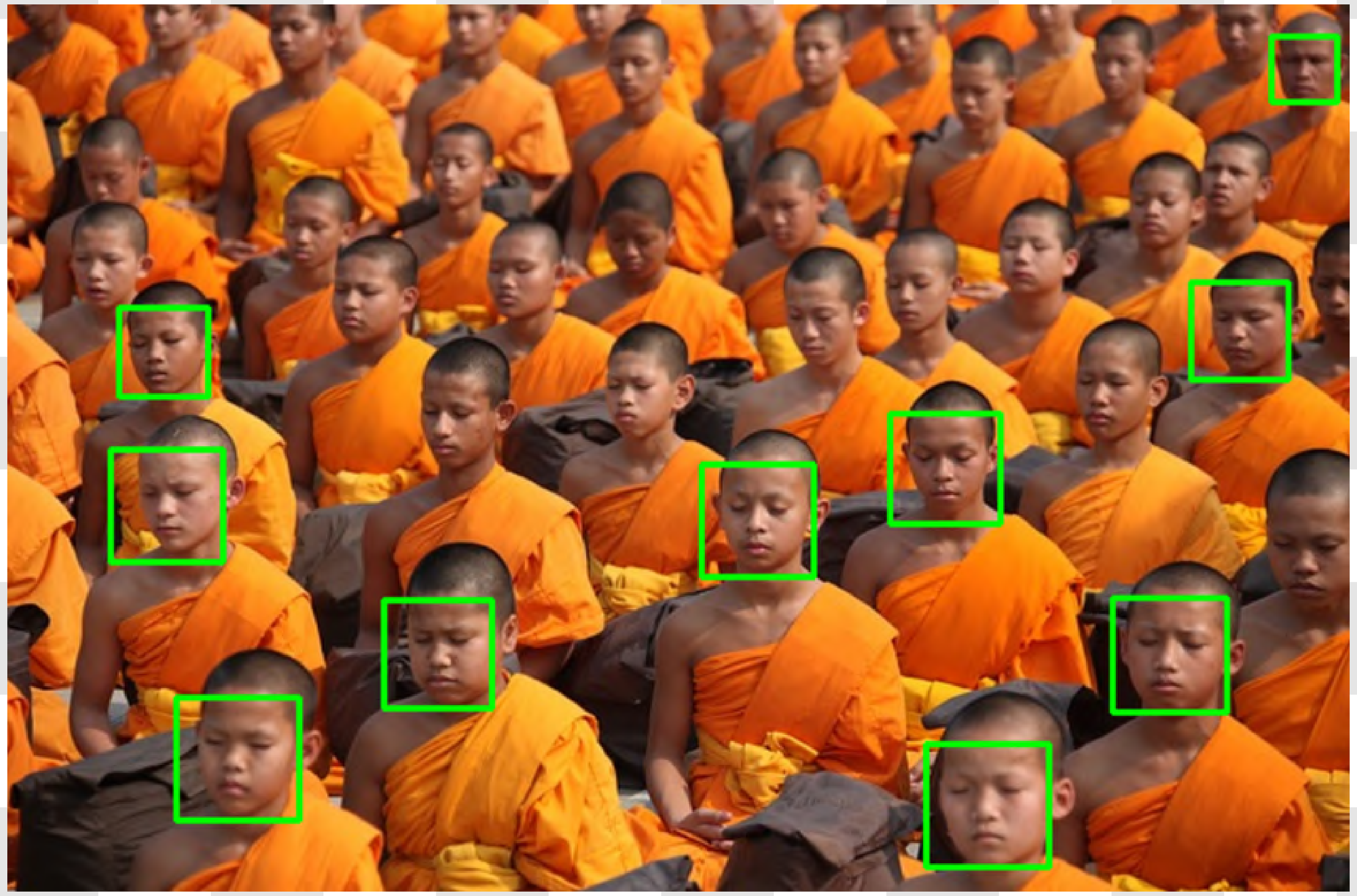
DNN Caffe



HoG



DNN YOLO



Haar Cascade

DNN Caffe

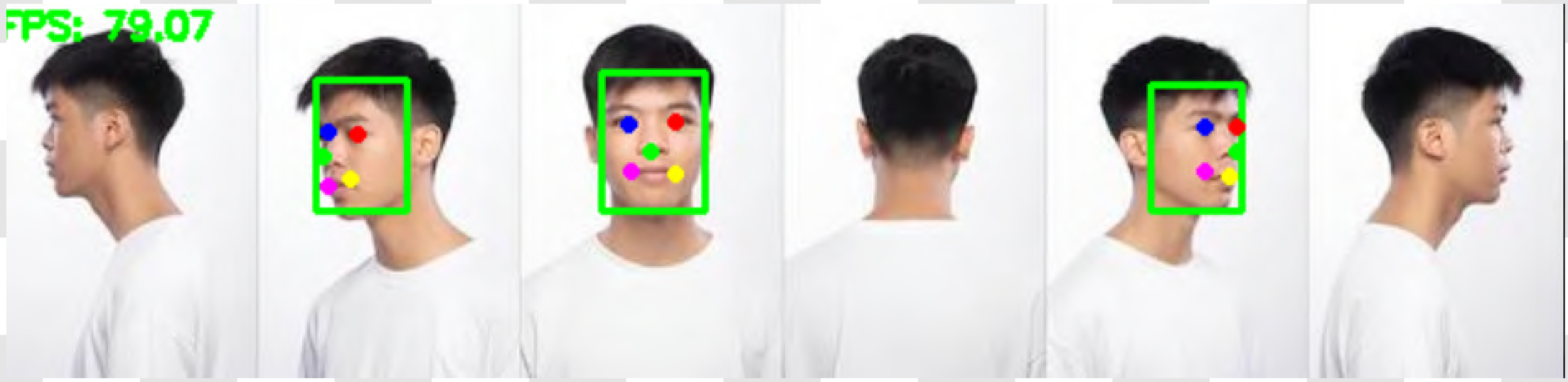
HoG

DNN YOLO

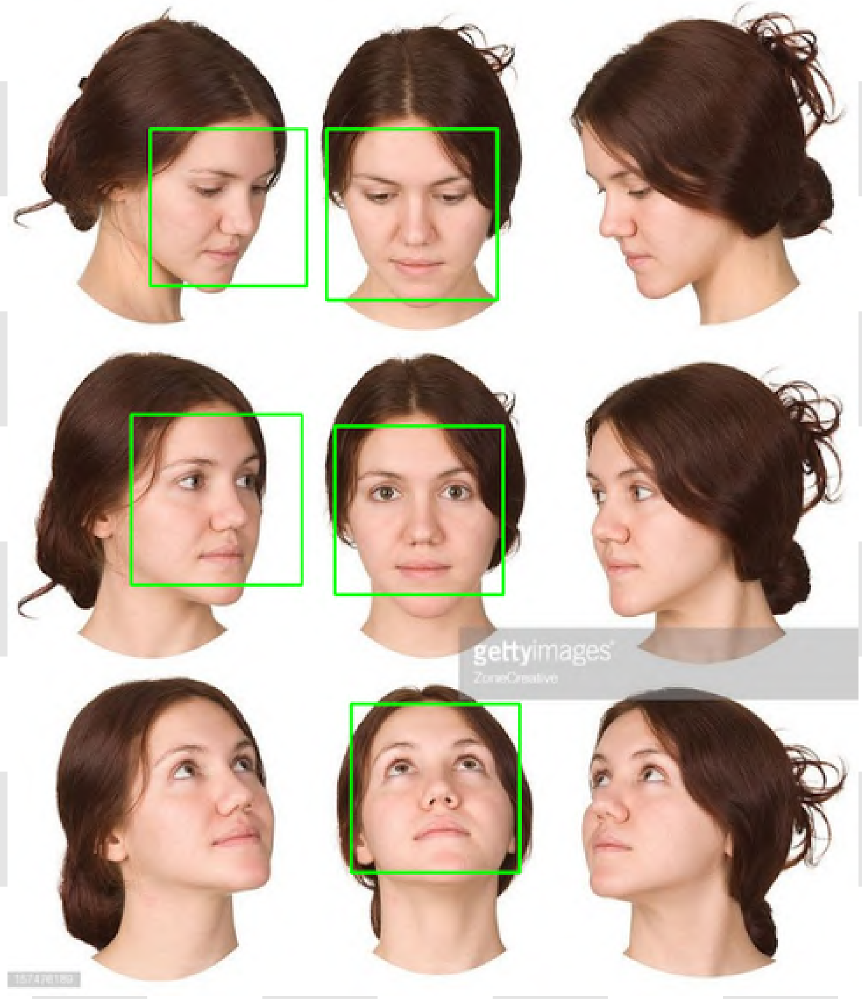
Haar Cascade

DNN Caffe

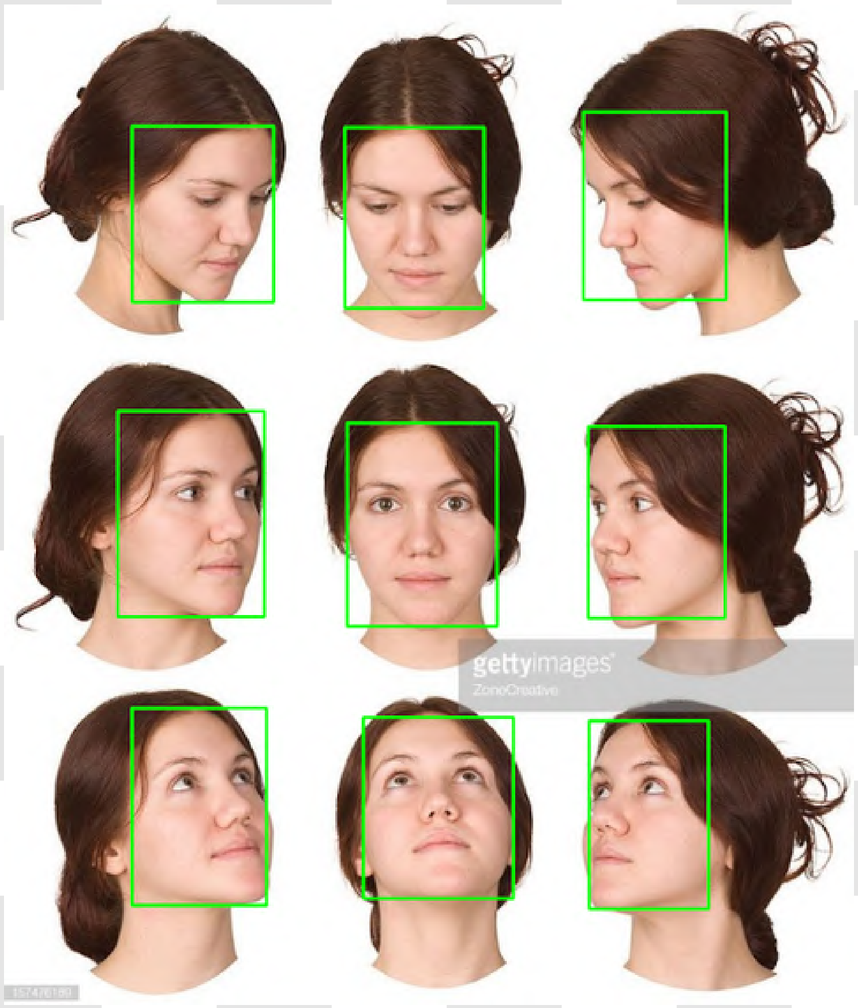
HoG



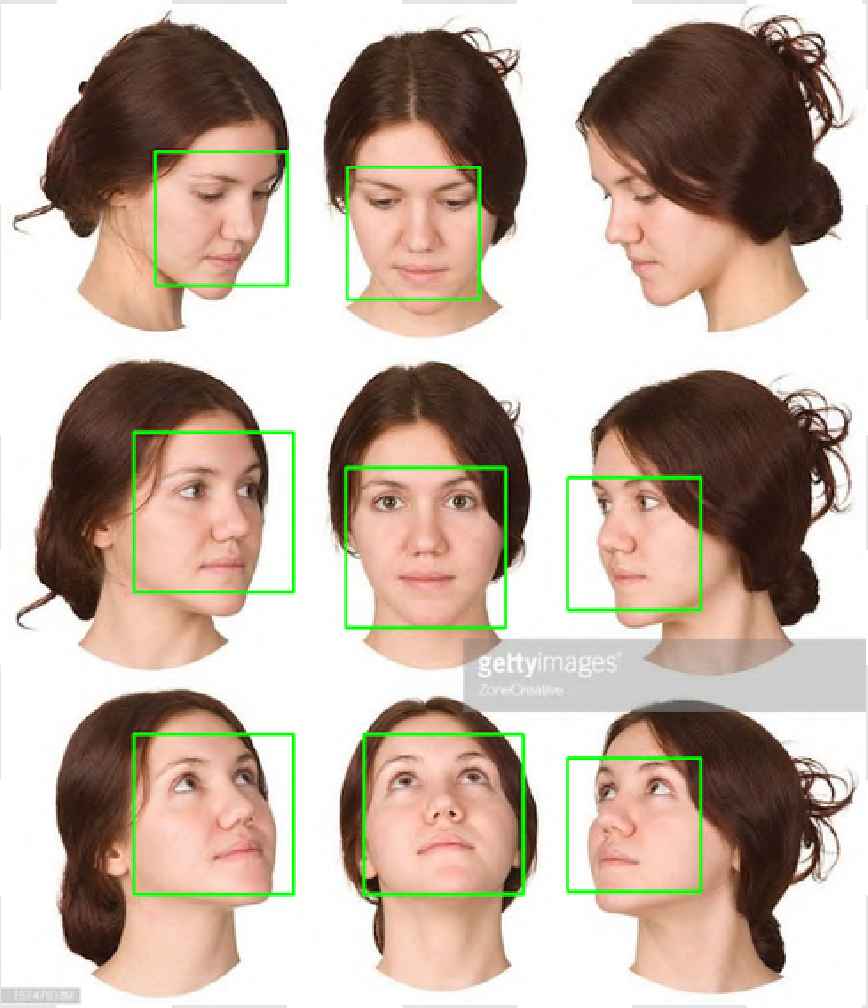
DNN YOLO



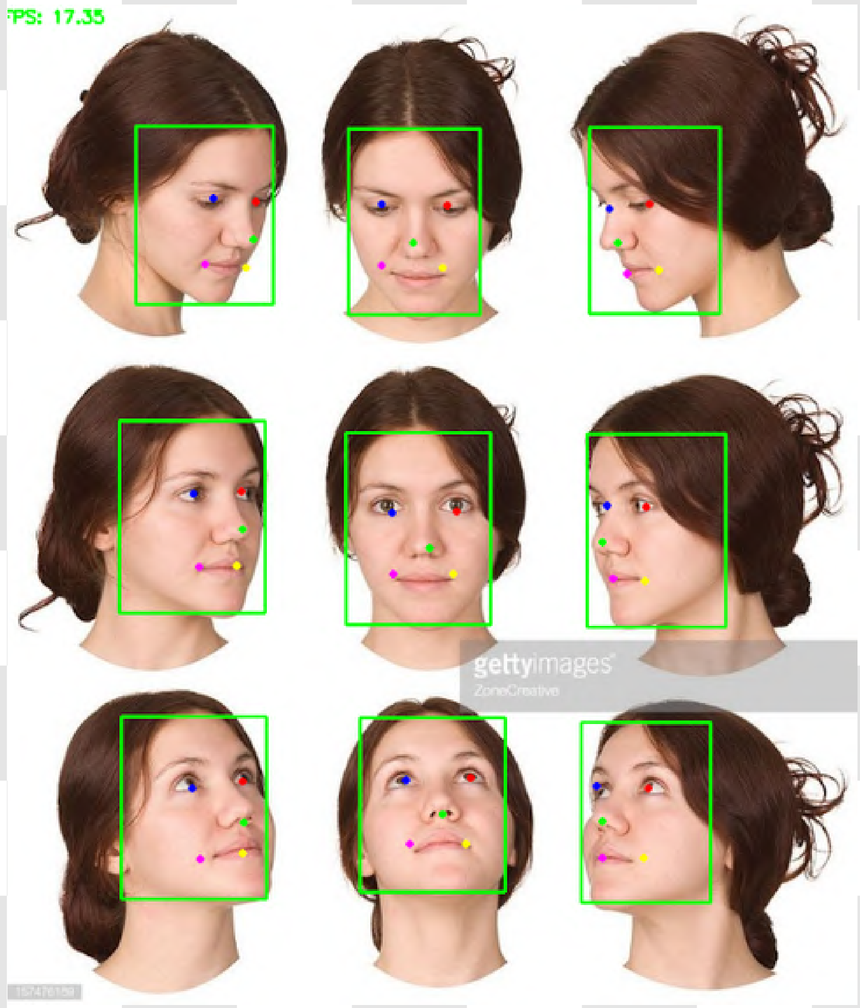
Haar Cascade



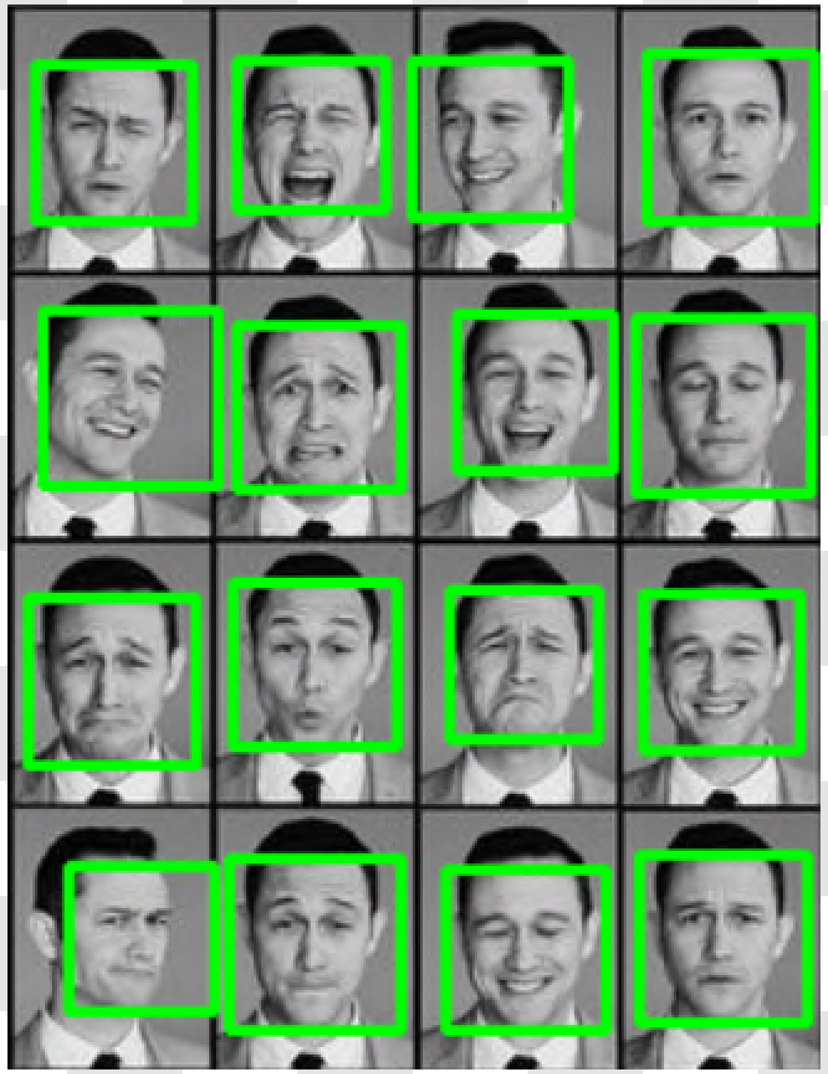
DNN Caffe



HoG



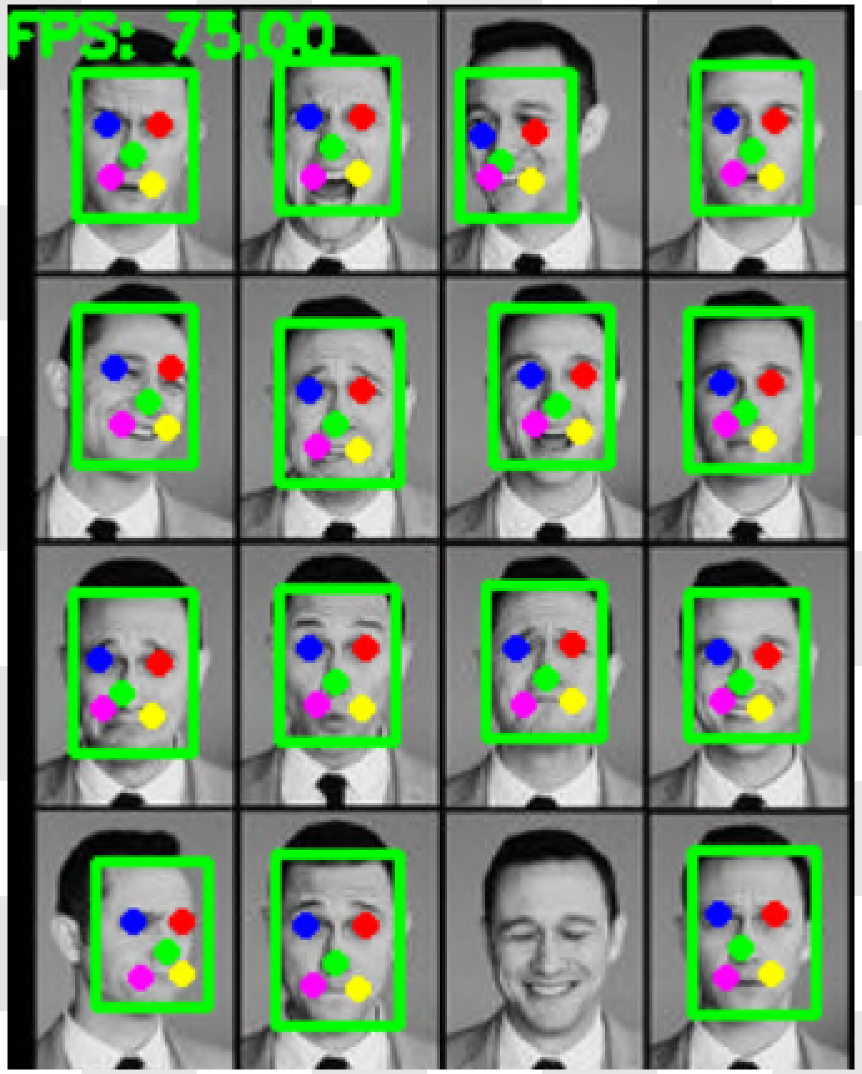
DNN YOLO



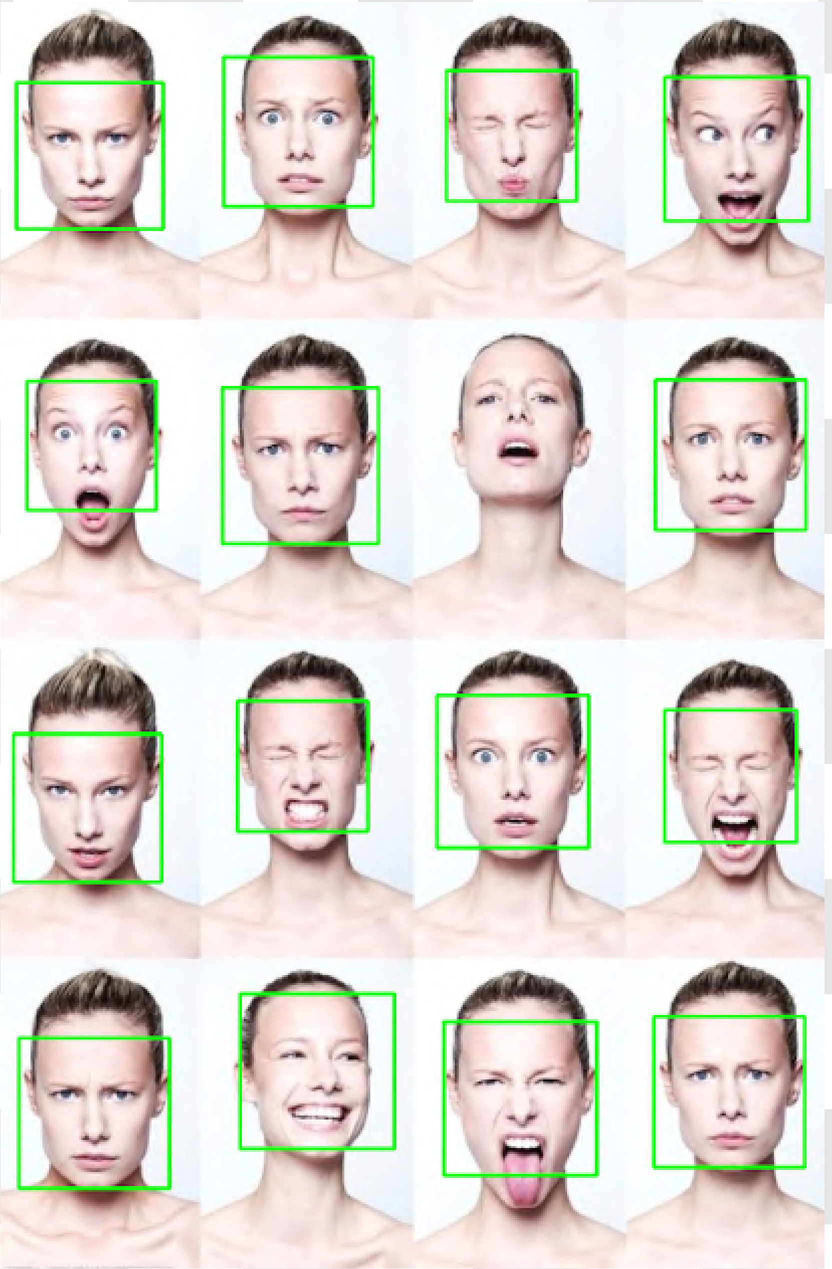
Haar Cascade

DNN Caffe

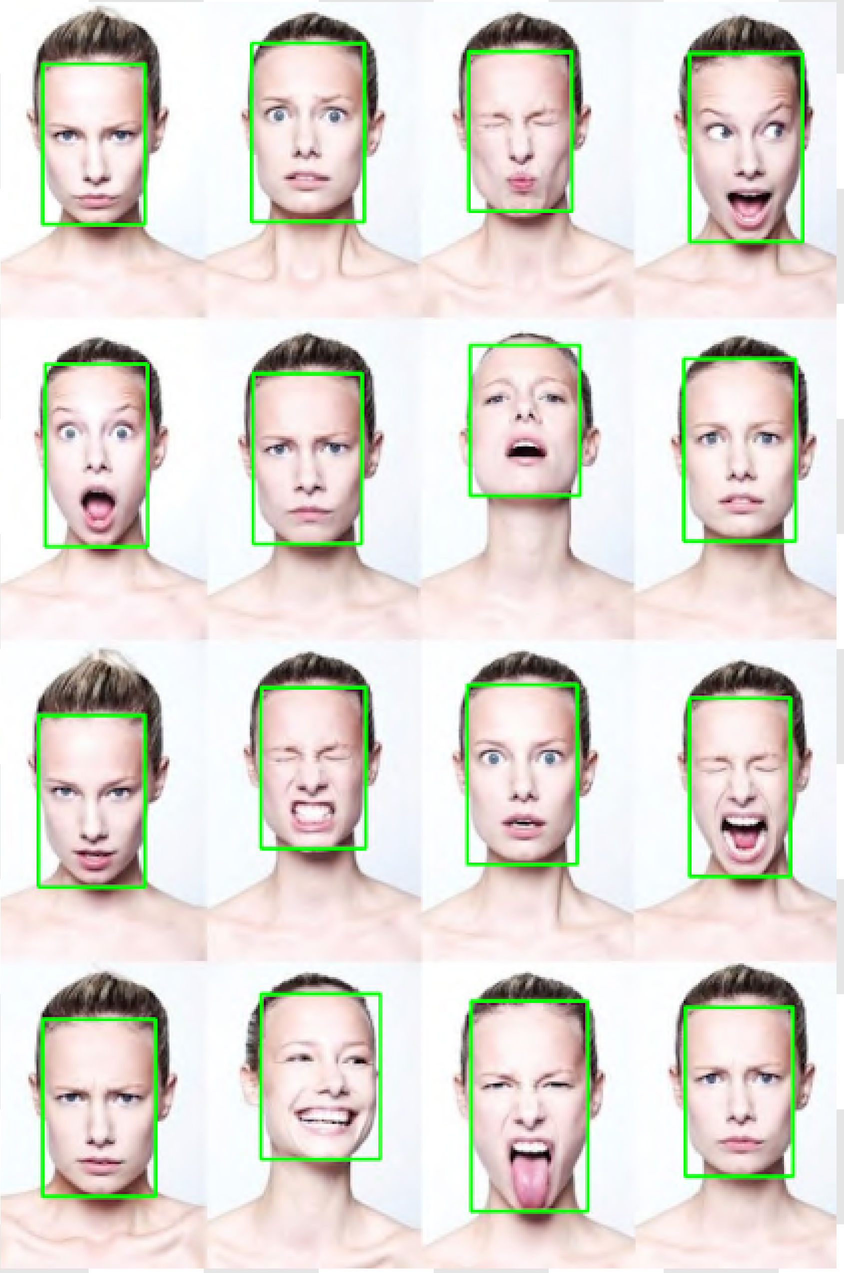
HoG



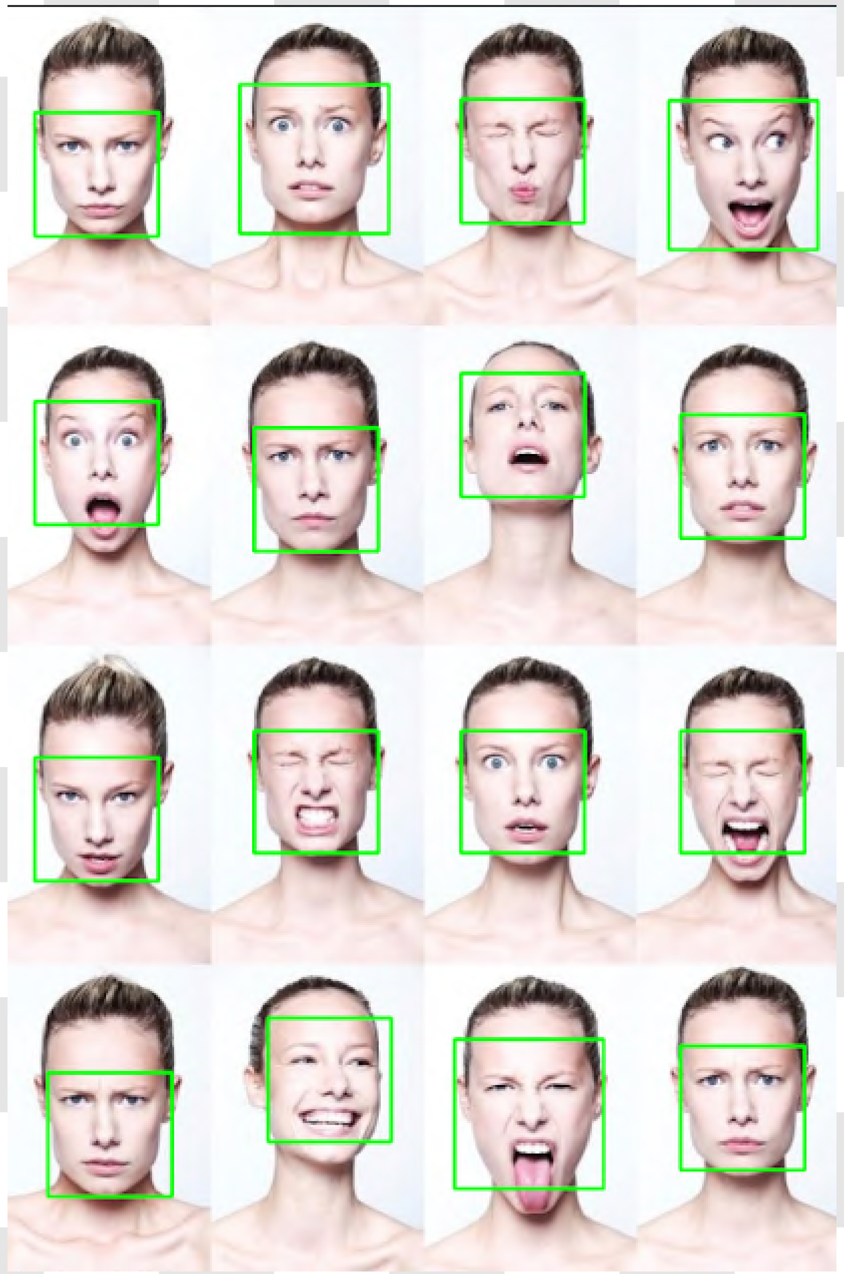
DNN YOLO



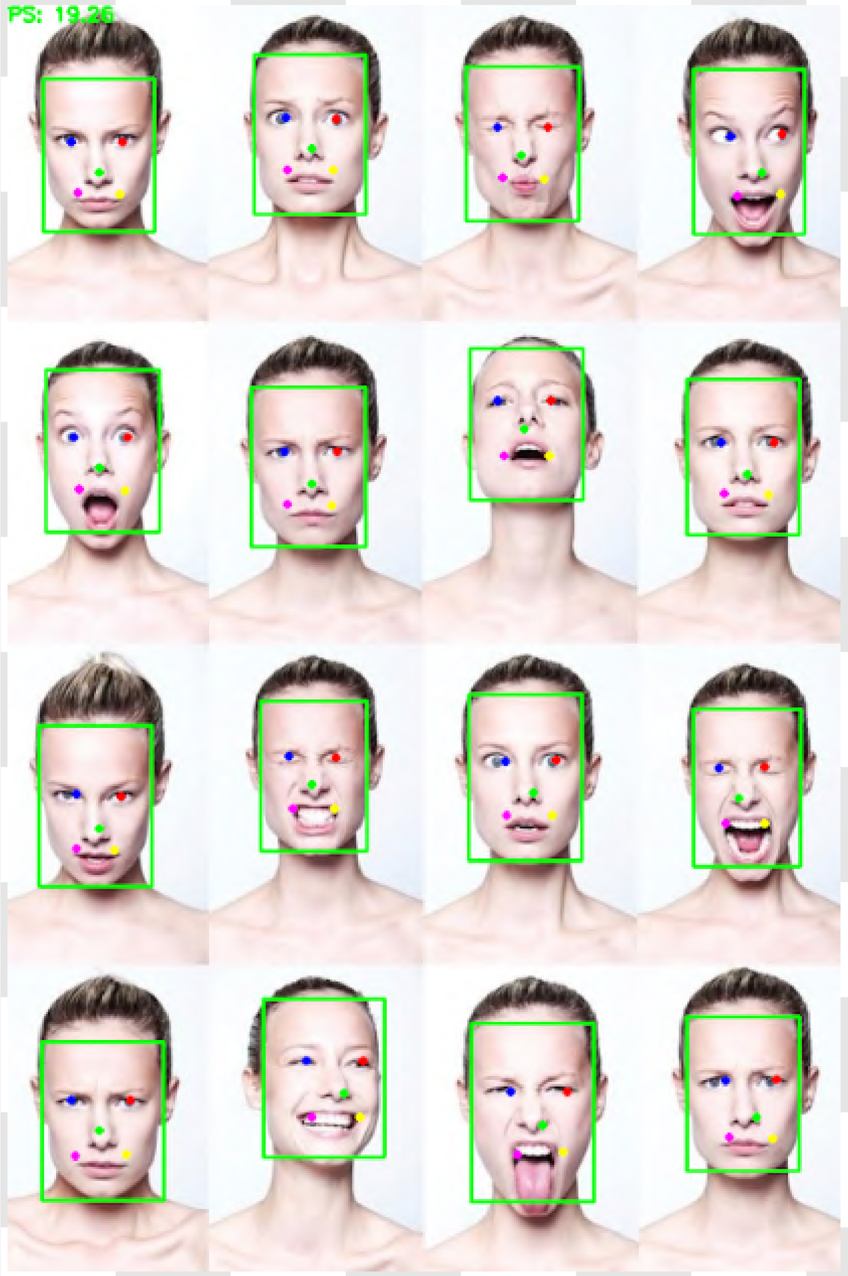
Haar Cascade



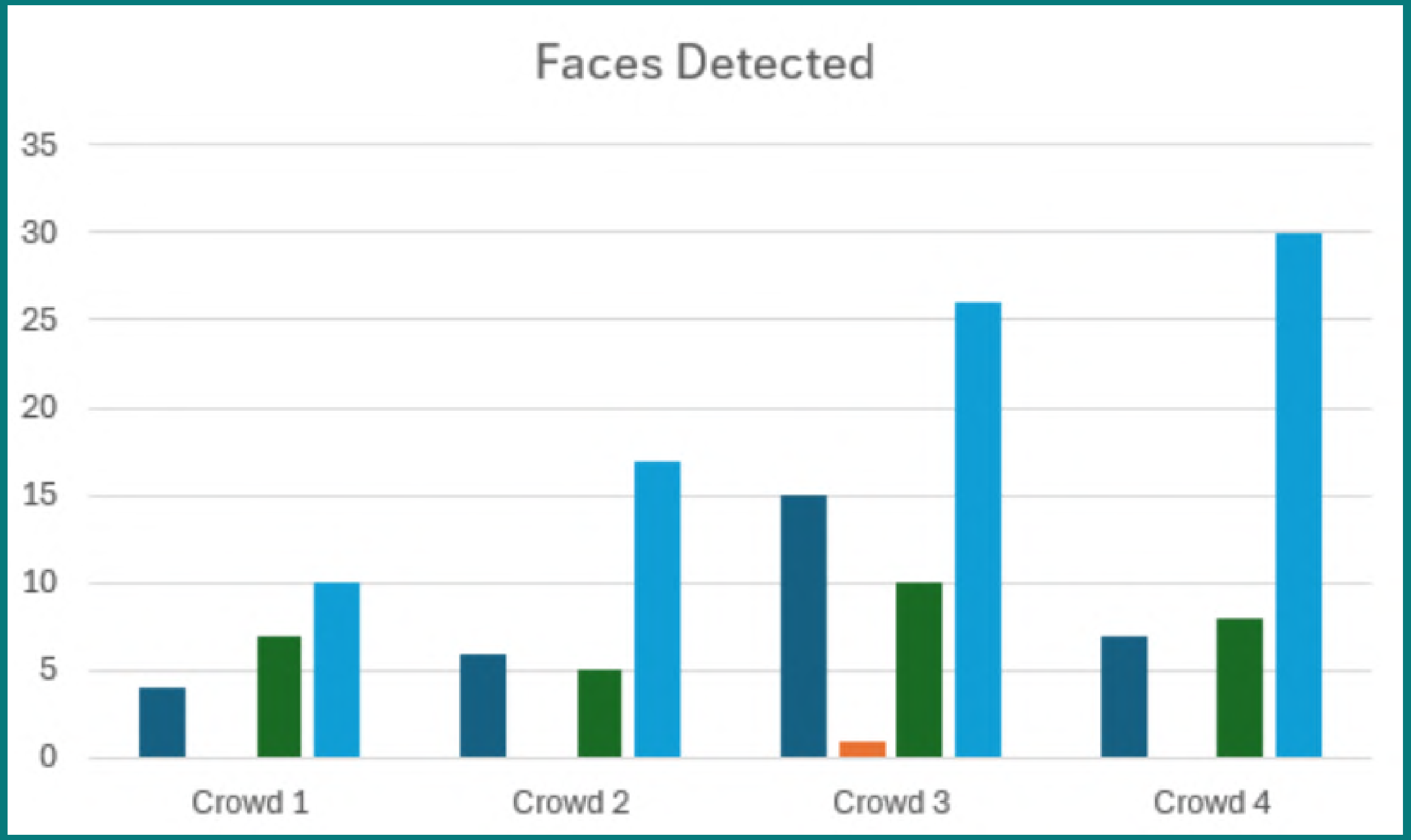
DNN Caffe

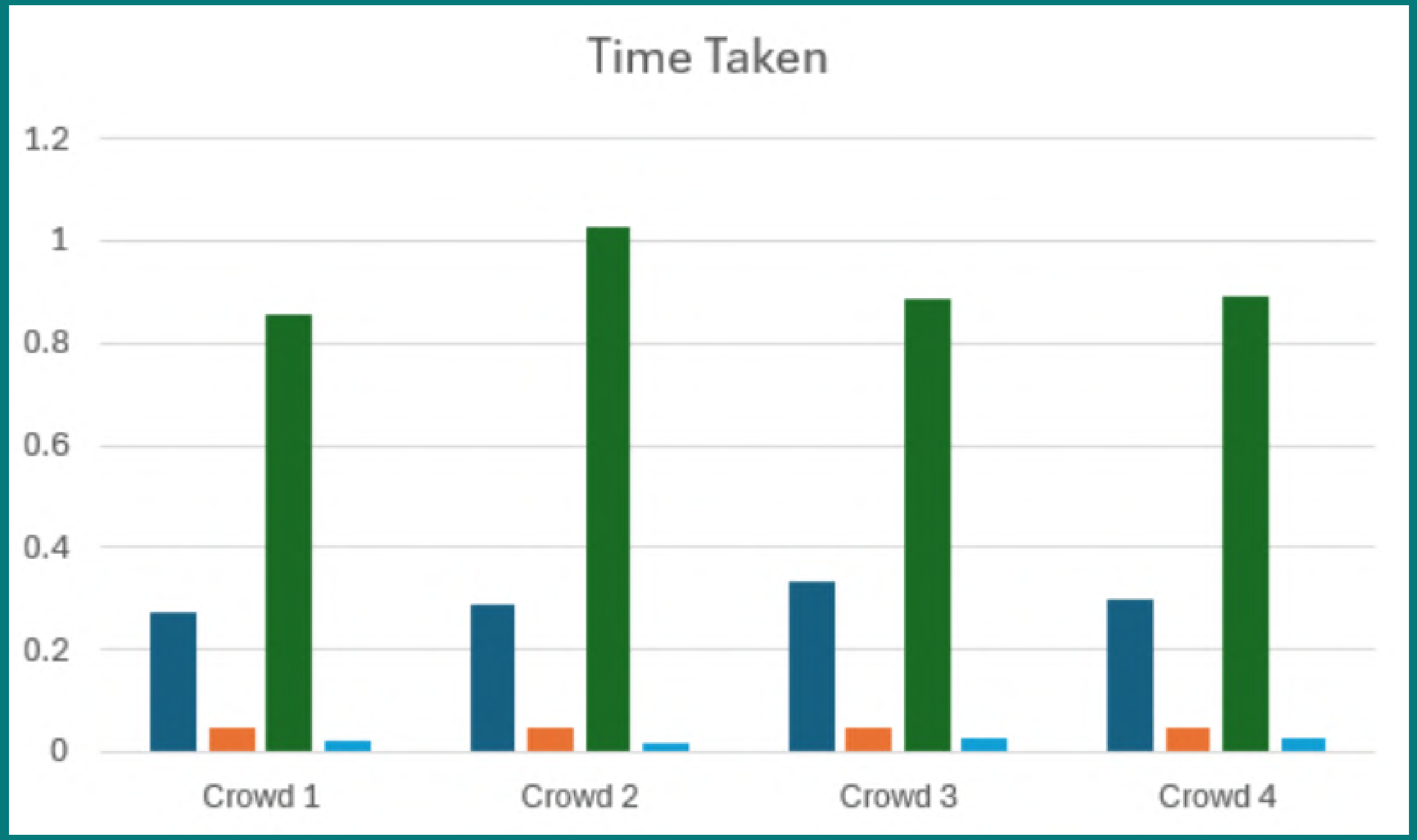


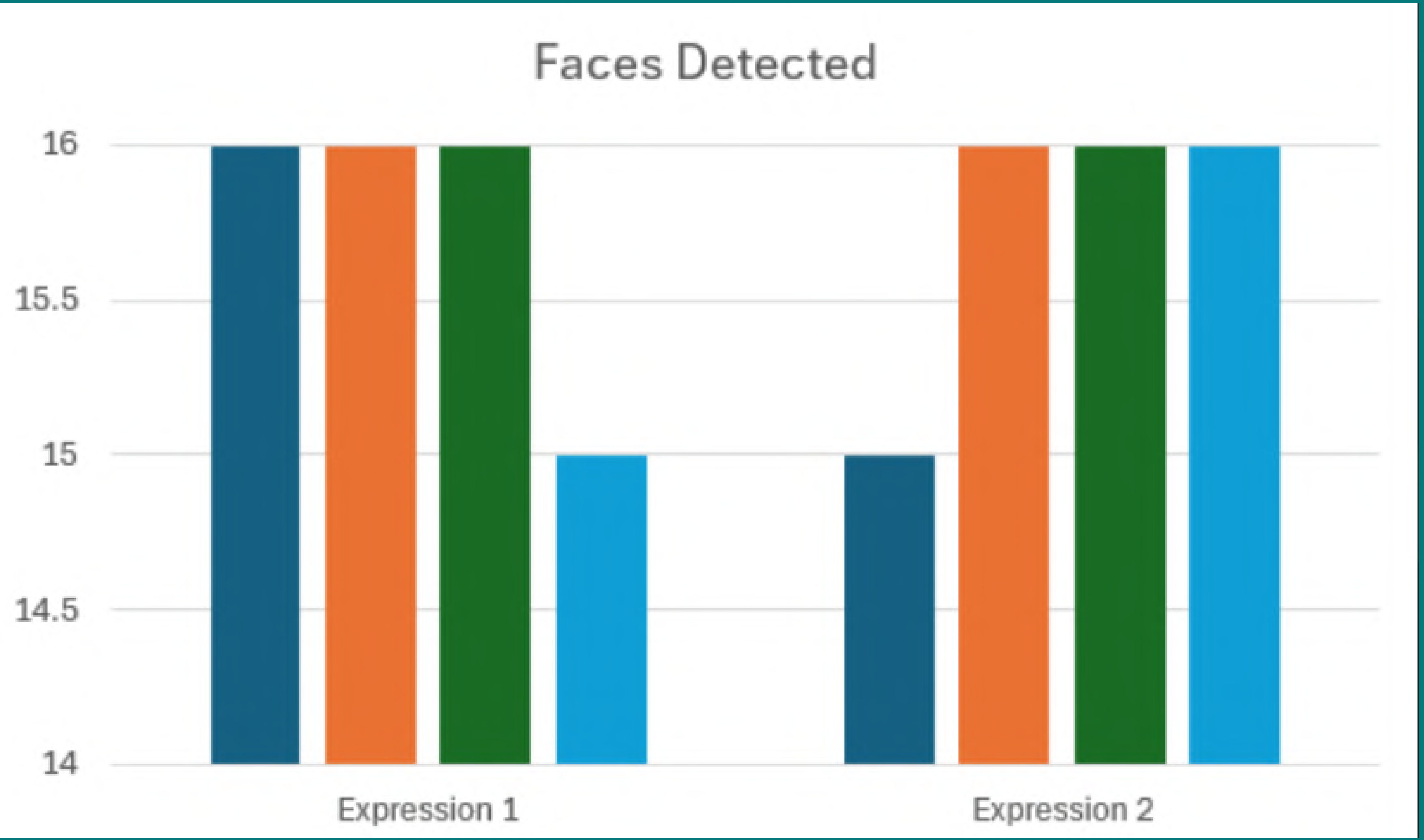
HoG

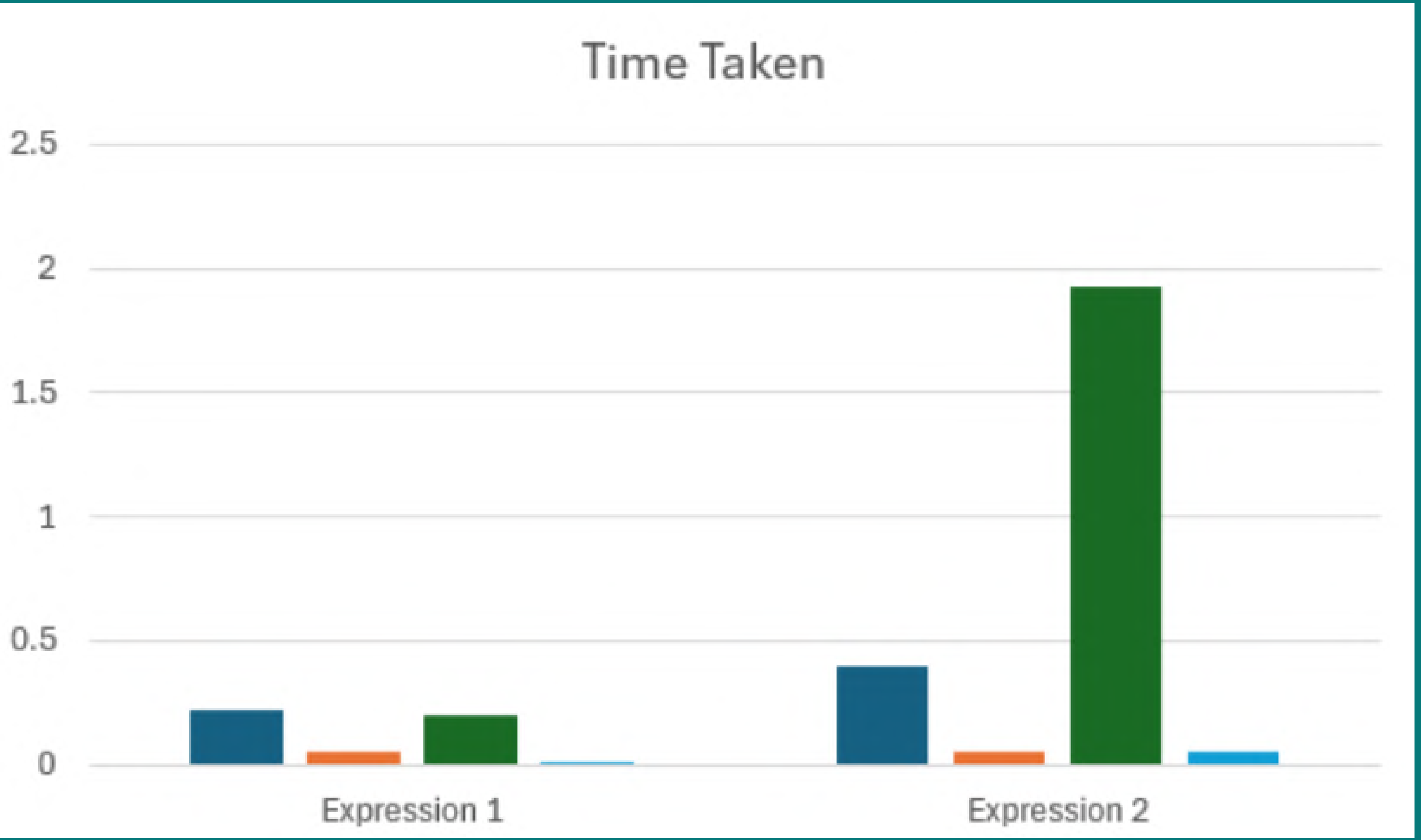


DNN YOLO

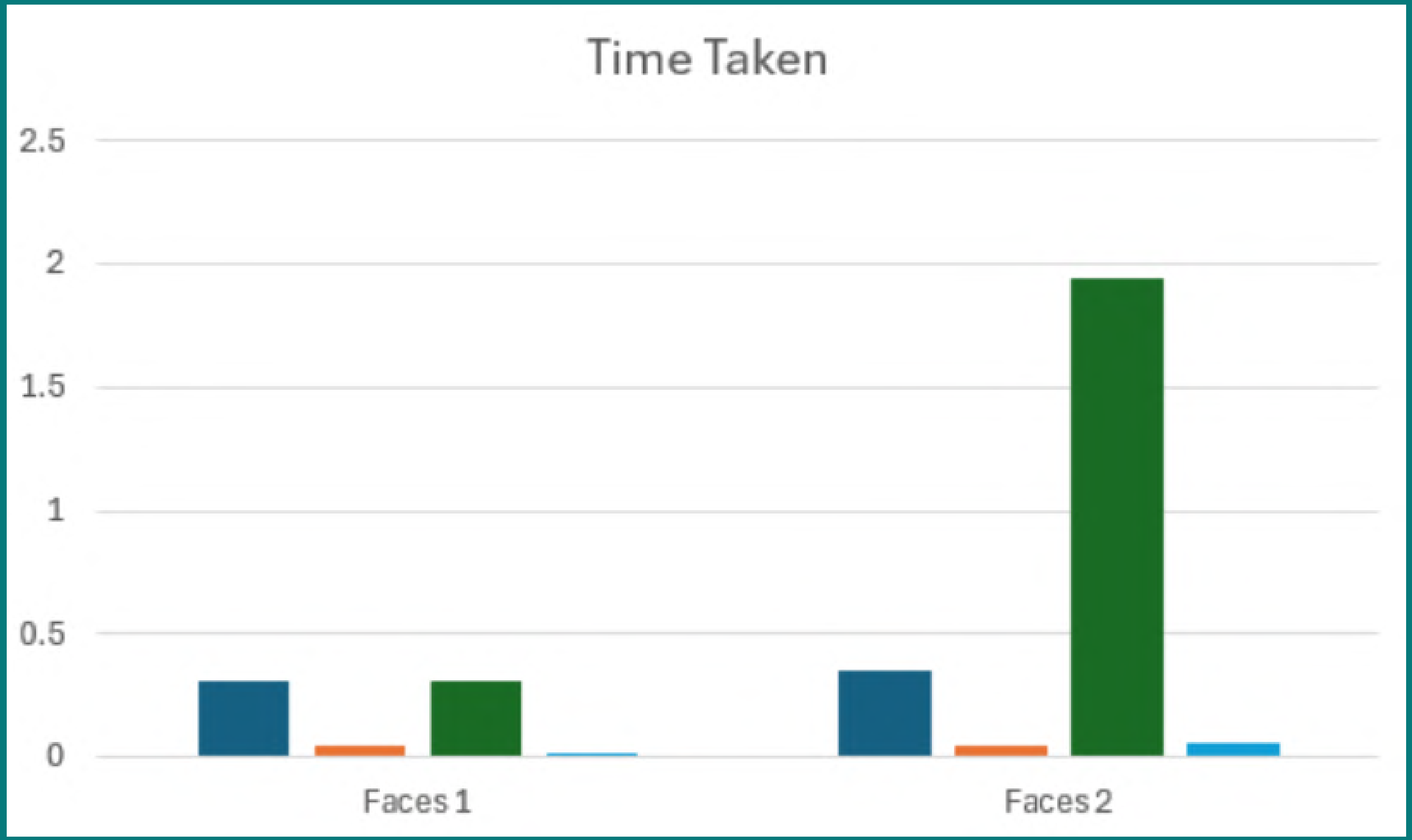

Faces Detected
35
30
25
20
15
10
5
0
Crowd 1
Crowd 2
Crowd 3
Crowd 4


Time Taken
1.2
1
0.8
0.6
0.4
0.2
0
Crowd 1
Crowd 2
Crowd 3
Crowd 4

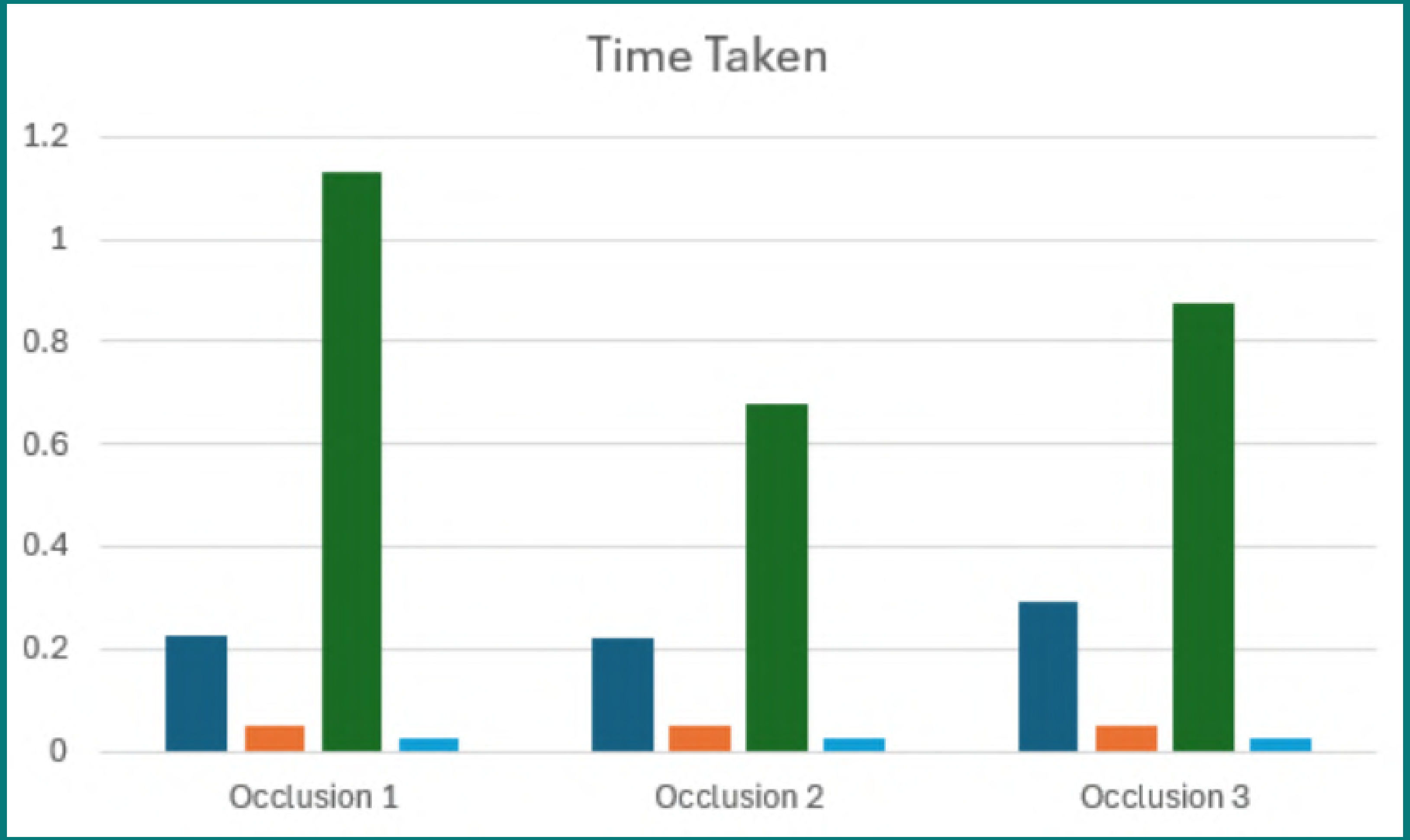




Back to Agenda Page



Back to Agenda Page



Back to Agenda Page

	Haar Cascade		DNN-Caffe		HoG		DNN-YOLO	
	Faces detected	Time Taken	Faces detected	Time Taken	Faces detected	Time Taken	Faces detected	Time Taken
Crowd 1	4	0.275	6	0.05	15	0.858	7	0.0238
Crowd 2	0	0.2873	0	0.05	1	1.03	0	0.02
Crowd 3	7	0.3322	5	0.05	10	0.888	8	0.0269
Crowd 4	10	0.3004	17	0.05	26	0.89	30	0.02707
Faces 1	1	0.3146	3	0.05	3	0.314	3	0.0126
Faces 2	5	0.3542	9	0.05	8	1.942	9	0.0576
Occlusion 1	0	0.2275	1	0.05	0	1.13	0	0.0234
Occlusion 2	0	0.223	4	0.05	0	0.679	0	0.0256
Occlusion 3	0	0.2898	1	0.05	0	0.873	0	0.0267
Expression 1	16	0.223	16	0.06	16	0.204	15	0.01333
Expression 2	15	0.3978	16	0.05	16	1.928	16	0.05193

Back to Agenda Page

FINDINGS

Back to Agenda Page

Accuracy

It is detected by upscaling the images for HoG by a factor of 2 to ensure better results. Haar is the most inaccurate of the lot. Both the CNN methods do well in terms of accuracy and HoG is also quite accurate when upscaled.

Upscaling improves HoG accuracy.

Back to Agenda Page

Speed

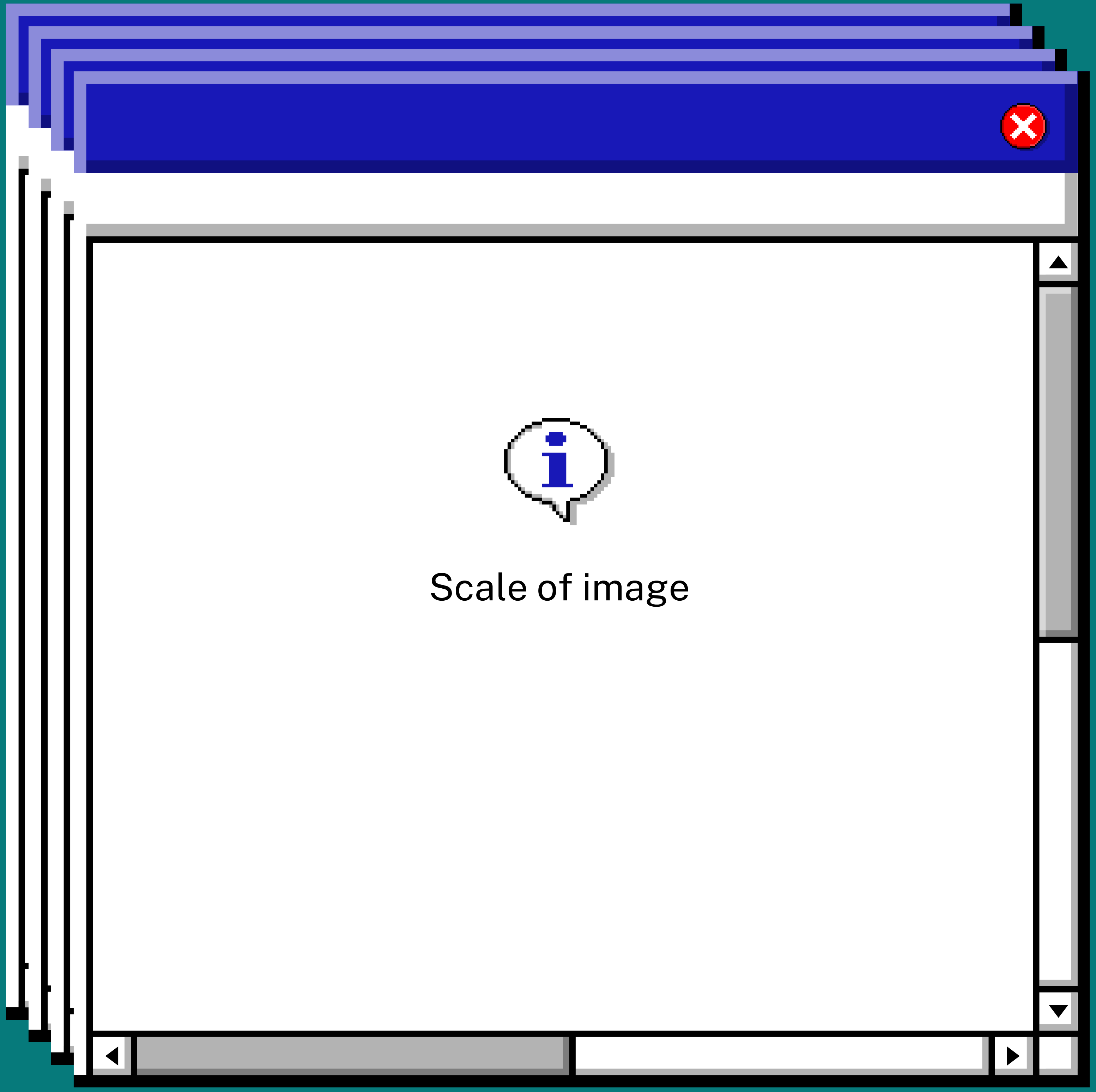
HoG takes a longer time due to the upscaling required. Both the CNN methods outperform the other two in terms of speed. DNN-YOLO method however is faster as it performs object detection in a single pass through the neural network instead of sliding windows or region based approaches.

HoG is slower due to upscaling.

Back to Agenda Page

Varying Conditions

HoG can detect faces of size up to ~ (70×70) after which they fail to detect. This is the major drawback as it is impossible to know the size of the face beforehand in most cases, we can get rid of this problem by upscaling the image, but then the speed advantage as compared to DNN goes away.



Back to Agenda Page

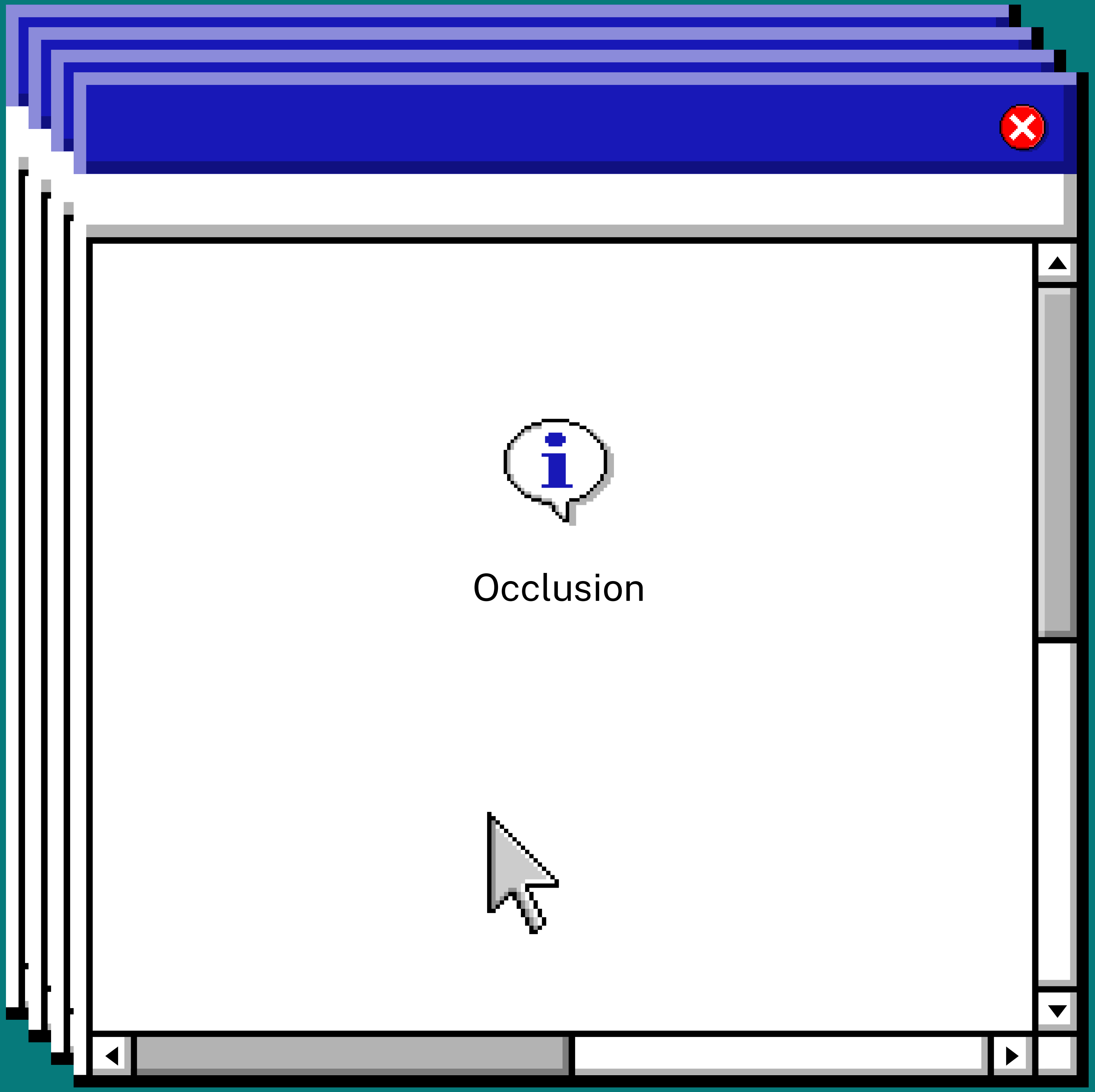
Varying Conditions

As expected, Haar based detector fails. The HoG-based detector does detect faces for left or right-looking faces (since it was trained on them) but is not as accurate as the DNN-based detectors of OpenCV and Dlib.

Back to Agenda Page

Varying Conditions

The DNN Caffe method outperforms the others, even though it is slightly inaccurate. This is mainly because the CNN features are much more robust than HoG or Haar features. The robustness of HoG depends on the scale of the images. It performs well with upscaled images but fails when faces are smaller.

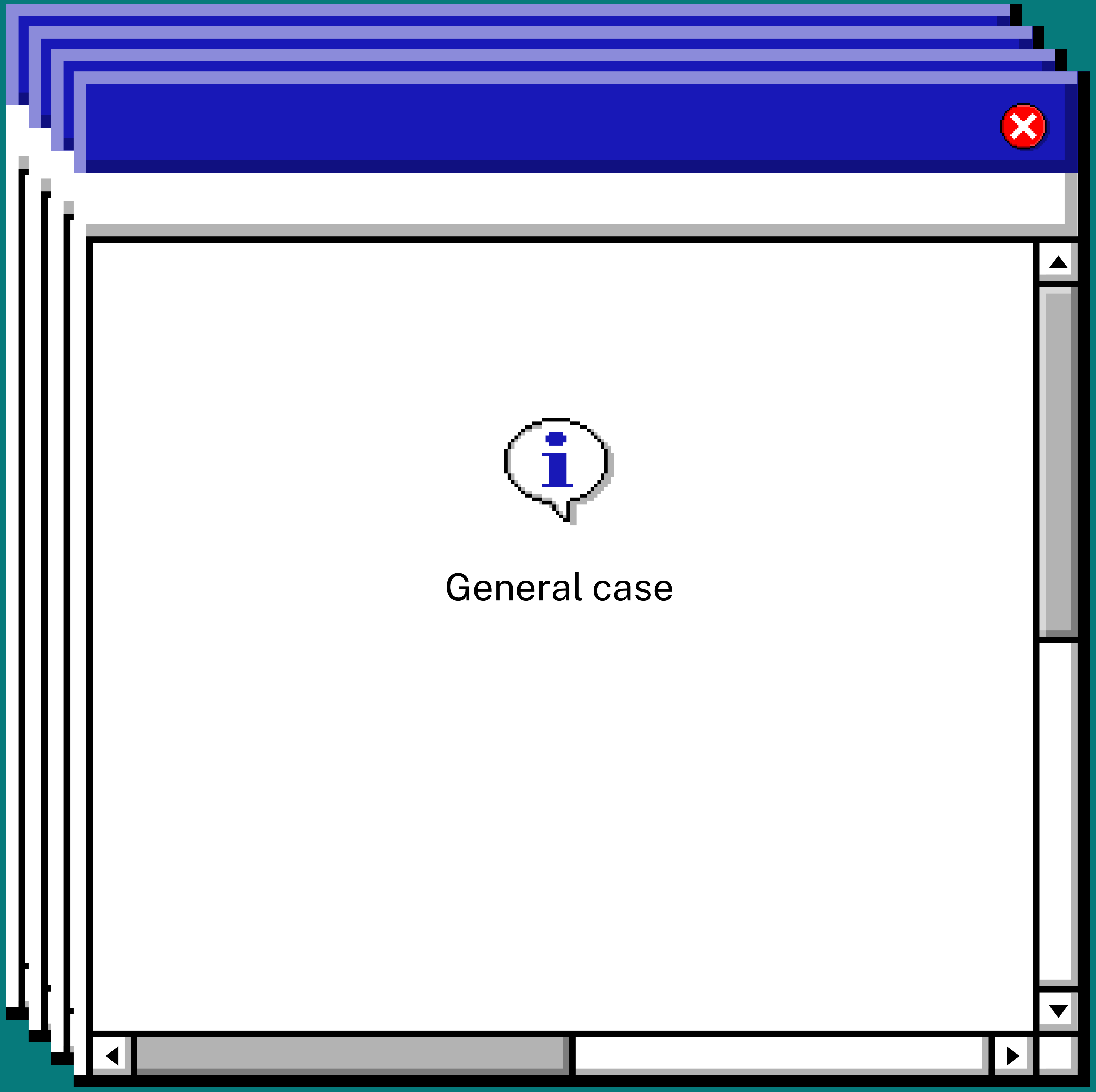


Back to Agenda Page

CONCLUSION AND FUTURE SCOPE

Back to Agenda Page

In most applications, we won't know the face size in the image beforehand. Thus, it is better to use DNN-Caffe method as it is pretty-fast and very accurate, even for small sized faces. It also detects faces at various angles. We recommend using DNN-Caffe in most cases. However if your computer can handle high GPU Computations then we recommend DNN-YOLO.



Back to Agenda Page

Dlib HoG is the fastest method on the CPU. But it does not detect small sized faces (< 70x70). So, if you know that your application will not be dealing with very small sized faces then HoG based Face detector is a better option.

Medium to large image sizes

Back to Agenda Page

Since feeding high-resolution images is not possible with these algorithms (for computation speed), HoG detectors might fail when you scale down the image. On the other hand, the DNN-YOLO method can be used for these since it detects small faces.

High-resolution images

Back to Agenda Page

Haar Cascade can be used in scenarios that require real-time applications with limited computational resources.

DNN-Caffe can be used in applications where accuracy is critical and computational resources are sufficient.

HoG Model are preferred when accuracy and speed is important but computational resources are limited.

DNN YOLO Model excels in applications requiring both accuracy and speed and where computational resources are available.

Comparison

Back to Agenda Page

Future Scope

Investigating advanced techniques to enhance the speed and accuracy of face detection algorithms, potentially leveraging parallel processing or optimized architectures.

Exploring the integration of machine learning methods for adaptive face detection, enabling models to learn and adapt to new environments and challenges autonomously.

Extending the evaluation to larger and more diverse datasets to further validate the performance and robustness of the algorithms across various real-world scenarios.

Back to Agenda Page

References

[1] OpenCV, "Cascade Classifier," OpenCV Documentation, 3.4, 2018. [Online]. Available: https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html.

[2] M. Fabien, "Face Detection," Maël Fabien Tutorials, [Online]. Available: https://maelfabien.github.io/tutorials/face-detection/#2-imports.

[3] LearnOpenCV, "Object Detection using YOLOv5 and OpenCV DNN in C++ and Python," LearnOpenCV, [Online]. Available: https://learnopencv.com/object-detection-using-yolov5-and-opencv-dnn-in-c-and-python/.

[4] OpenCV, "YOLO Object Detection," OpenCV Documentation, 4.x, [Online]. Available: https://docs.opencv.org/4.x/da/d9d/tutorial_dnn_yolo.html.

[5] Unsplash. [Online]. Available: https://unsplash,com/.