**Visual Attendance System**

**1. Introduction**

In today's fast-paced world, attendance tracking is a fundamental aspect of various institutions, such as schools, universities, and organisations. The conventional methods of manual attendance systems, involving sign-in sheets and physical presence verification, often prove to be time-consuming, inefficient, and prone to errors [1]. To address these limitations, the advent of technology has paved the way for automated attendance systems that leverage the power of visual recognition [2]. The focus of this paper is on the development and evaluation of a cutting-edge technology known as the Visual Attendance System (VAS), which employs computer vision techniques to streamline the attendance management process.

The Visual Attendance System utilises state-of-the-art computer vision algorithms, including face detection, recognition, and tracking, to accurately identify and record the attendance of individuals within a given setting [4]. By harnessing the capabilities of advanced machine learning models and image processing techniques, VAS aims to revolutionise traditional attendance tracking systems by offering real-time, contactless, and highly accurate attendance monitoring [4].

This paper intends to provide an in-depth analysis of the Visual Attendance System, exploring its underlying principles, technical components, and potential applications. Furthermore, the study will evaluate the performance of VAS in terms of accuracy, efficiency, and scalability, comparing it with traditional attendance methods and other existing automated solutions. By examining the strengths, limitations, and future prospects of VAS, this research aims to contribute to the advancement of attendance management systems, with a particular focus on visual recognition technologies.

Overall, this paper aims to contribute to the growing body of knowledge on automated attendance systems by introducing the Visual Attendance System as a novel and promising solution. By combining cutting-edge computer vision techniques with efficient attendance management, VAS has the potential to enhance accuracy, reduce administrative burden, and facilitate a more streamlined and secure attendance monitoring process [3].

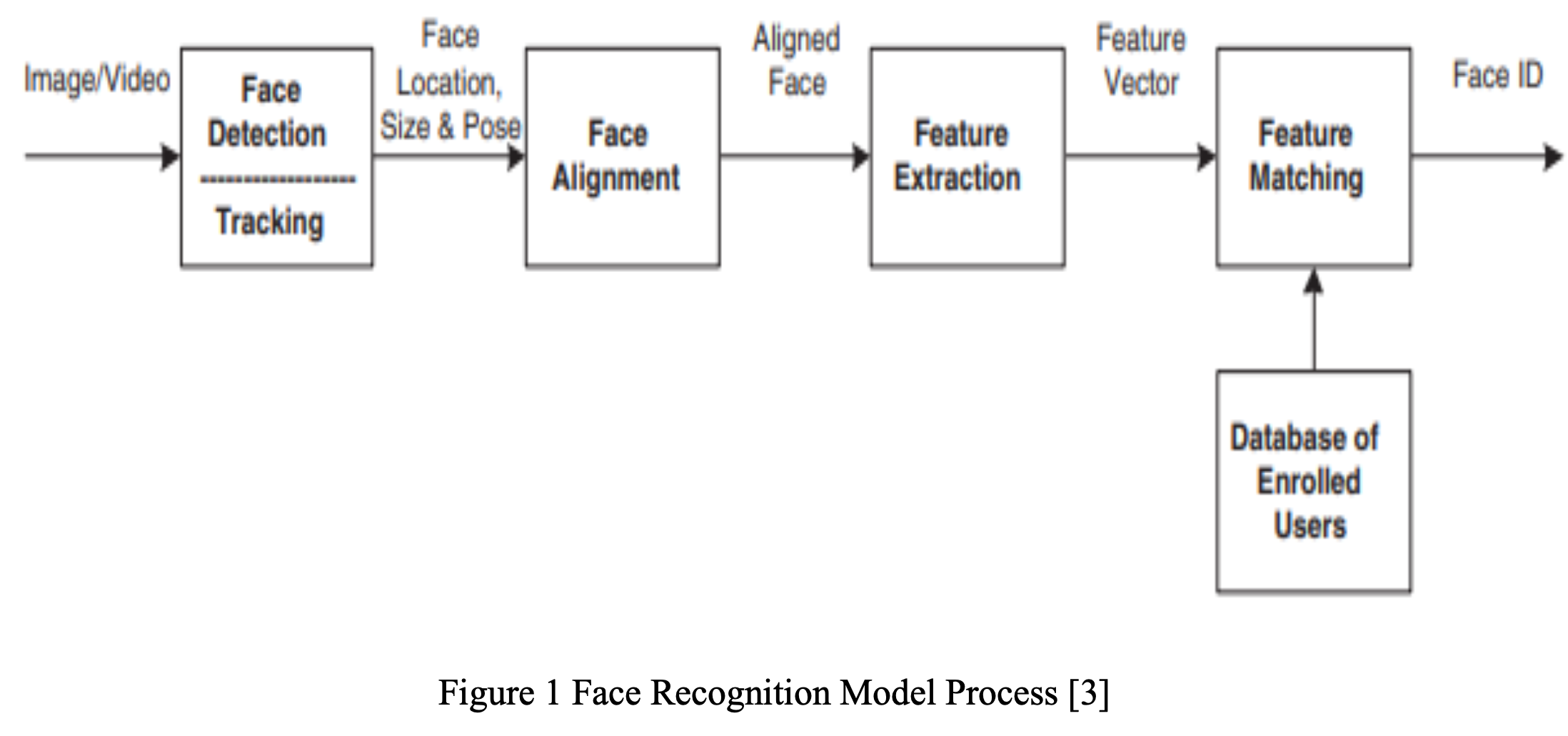


Figure 1 : Face Recognition Model Process [7]

**2. Literature**

1. *Face Detection*

To identify human faces in larger images and ignore other objects, face detection programs utilise algorithms and machine learning. Since human eyes are the easiest feature to identify, these algorithms often begin by searching for them. The system then proceeds to detect the mouth, nose, and eyebrows. The algorithm performs additional tests to confirm whether a face has been found. Improving the algorithm's ability to determine if there is a face in the image and its precise location requires training it on extensive datasets that include thousands of positive and negative photos [5][6].

There are two main types of face detection: static face detection and real-time face detection. In static face detection, a video or image is saved in the device's memory, providing the system with data for face detection. On the other hand, real-time face detection involves identifying people in real time from a live camera, such as a CCTV camera.

1. *Face Recognition*

Face recognition is a technology that goes beyond face detection. While face detection focuses on locating and identifying human faces in digital images or videos, face recognition takes it a step further by actually recognising and verifying the identity of specific individuals.

Face recognition systems use advanced algorithms and machine learning techniques to analyse and compare facial features, such as the size, shape, and positioning of the eyes, nose, mouth, and other facial landmarks. These unique facial characteristics are then matched against a database of known individuals to determine if there is a match.

C. *Face Detection Algorithms*

*C.1 Haar Cascade*

It is an object detection system that can recognise faces in still photos or moving videos. In order to recognise a face, Haar cascade classifiers are trained on a large number of pictures, both with and without faces. This means that instead of choosing non-facial areas, Haar classifiers scan the window for faces. According to research [6], image processing and pattern recognition experts have found it difficult to reliably identify human faces. An object detection system called the Haar Cascade Algorithm may be used to find faces, pedestrians, objects, and facial emotions [7].

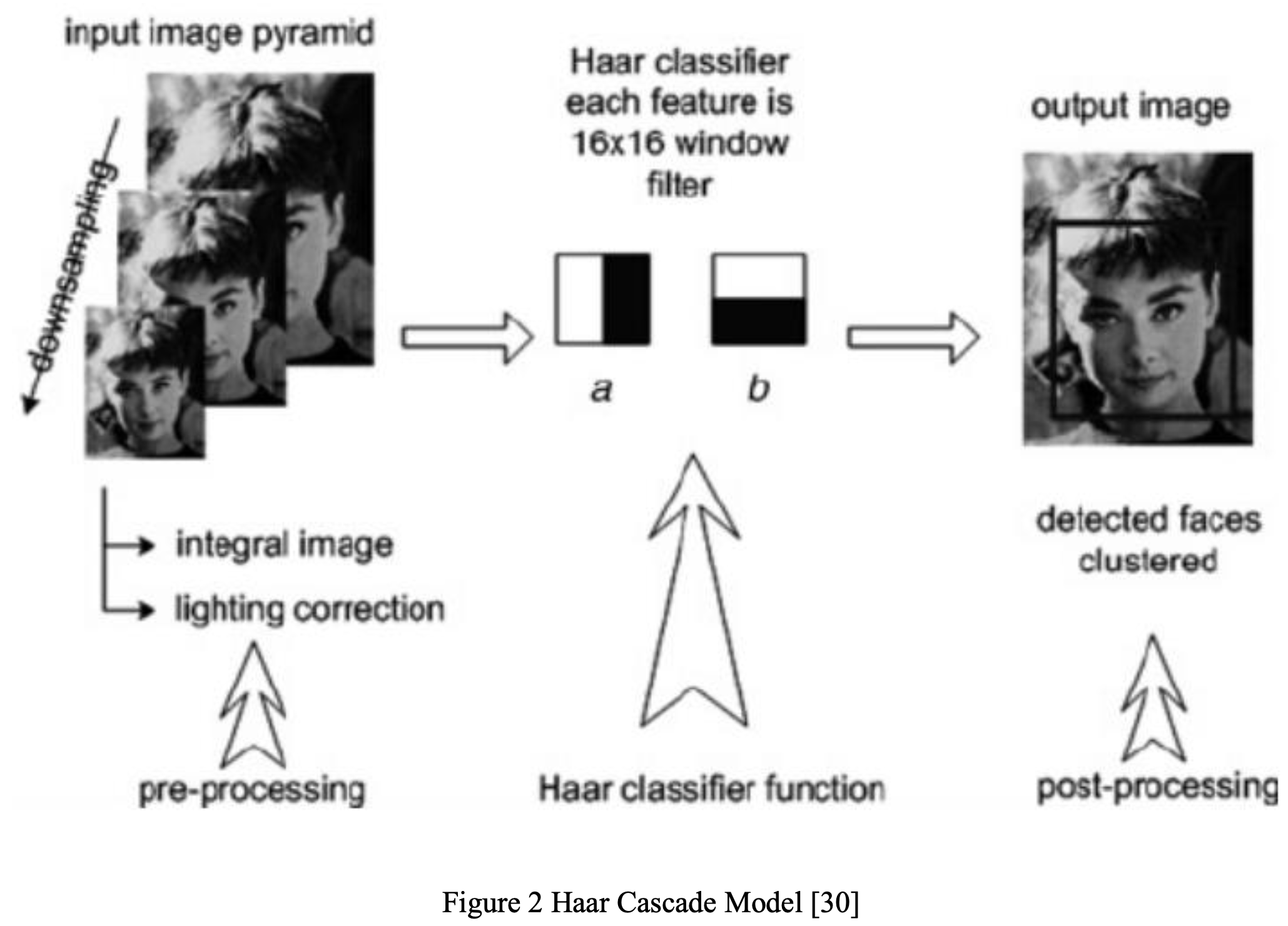


Figure 2 : Haar Cascade Model [11]

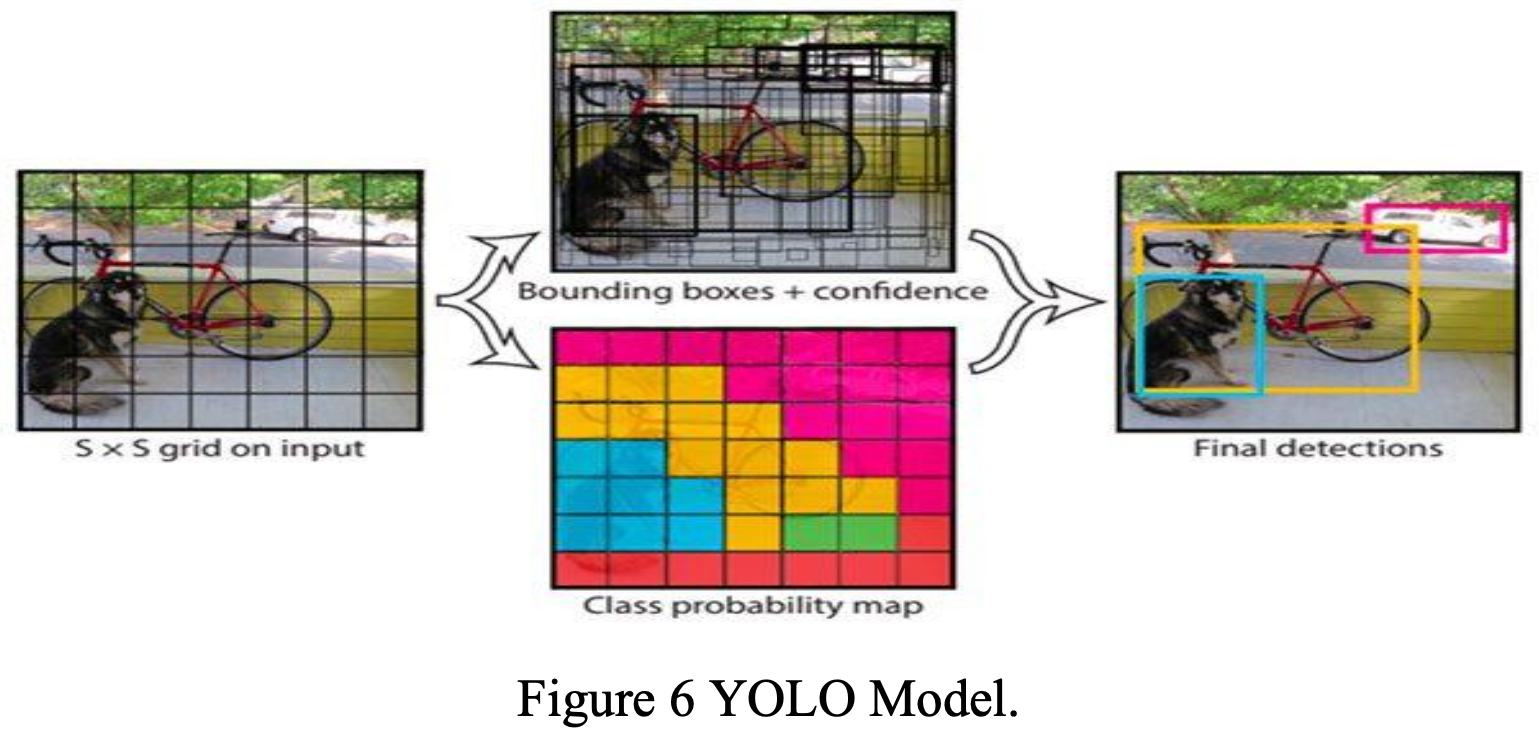


Figure 4 : YOLO Model

C.2 HoG (Histogram of Oriented Gradients) in Dlib library

The HoG feature descriptor uses histograms of gradient directions (oriented gradients) to represent features. Gradients, which are the x and y derivatives of an image, are particularly useful in capturing information about object shape in areas with edges and corners where there are sudden changes in intensity. These areas provide more valuable information than flat regions [8] [9].

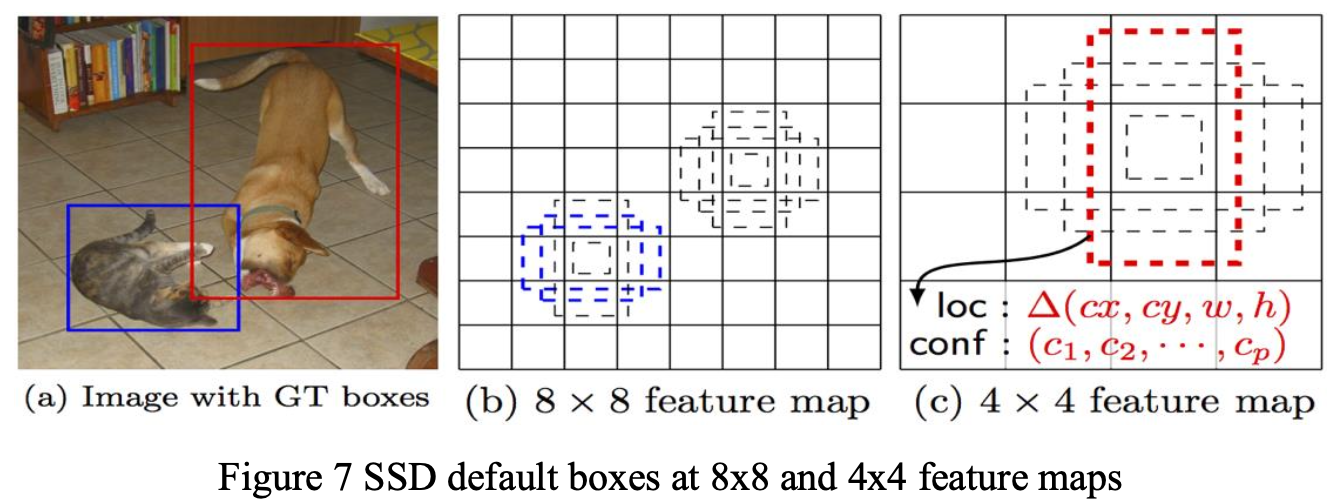


Figure 5 : SSD default boxes at 8x8 & 4x4 feature maps

Dlib, a software library, already has a pre-trained model for face identification. To use it, we simply need to create an object based on that model. This completes the process of predicting faces using the HOG feature descriptor and Dlib [10].

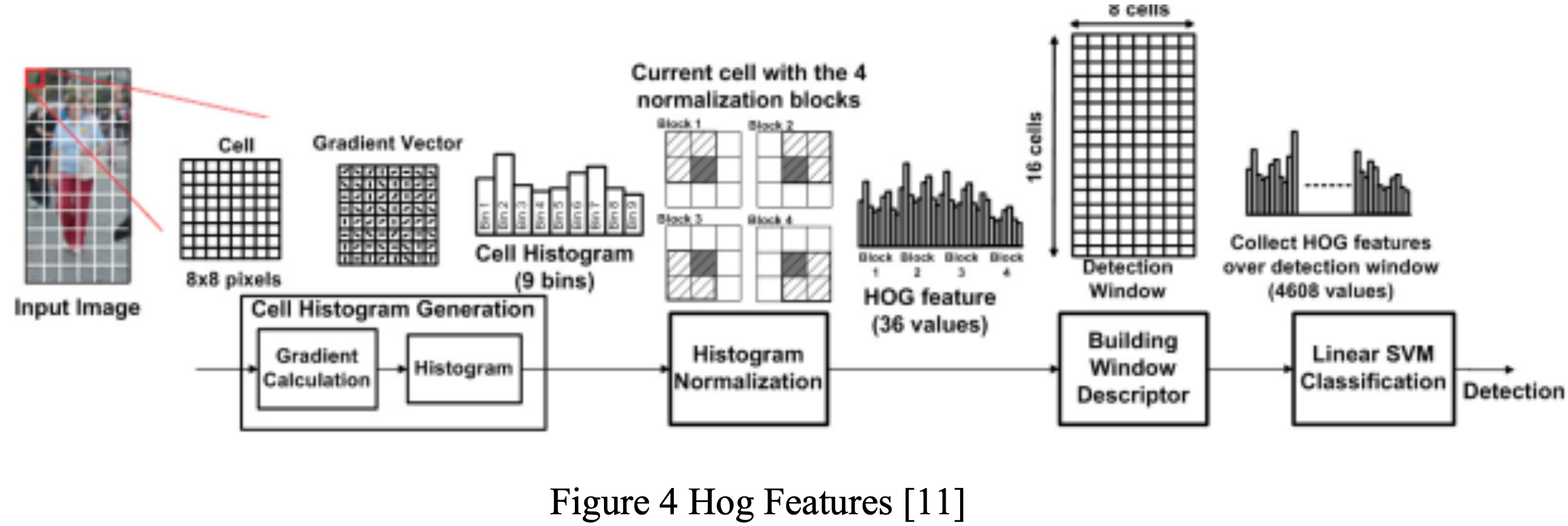


Figure 3 : HoG Features [12]

*C.3 YOLO*

One of the most advanced real-time object identification systems based on deep learning is called You Only Look Once (YOLO). It utilises a single neural network that analyses an image, divides it into regions, and calculates the probabilities of bounding boxes and the probabilities for each region. The YOLO model can process images in real-time at 45 frames per second (FPS). There is also a smaller version of the network called Fast YOLO, which achieves an impressive 155 FPS while still maintaining double the map (mean average precision) compared to other real-time detectors.

While YOLO has a lower tendency to predict false positives in the background compared to cutting-edge detection algorithms, it does have a higher rate of localisation errors [13].

Processing images with YOLO is simple and straightforward. The system,

Step 1) resizes the input image to 448 × 448,

Step 2) runs a single convolution network on the image

Step 3) thresholds the resulting detections by the model's confidence [13]

*C.4 Single Shot Multibox Detector (SSD)*

The network performs object localization and classification tasks in a single forward pass, which is known as a one-shot process. The method used for bounding box regression, introduced by Szegedy et al., is called MultiBox. The network functions as an object detector, identifying and classifying the objects it detects [13].

The design of SSD is inspired by the VGG-16 architecture, as shown in the figure above. However, it does not include fully connected layers. The VGG-16 network was selected for its impressive performance in image categorization tasks and its potential for transfer learning to improve results in various problems.

Instead of using the original VGG fully connected layers, SSD incorporates several auxiliary convolution layers starting from conv6. These additional layers enable feature extraction at multiple scales while progressively reducing the input size for each subsequent layer [13].

C.5 Similar Projects

While there are similar services to ours, our aim is to incorporate the best features from all of them into our own service. In the section below, we will compare and contrast these services, highlighting both their positive and negative characteristics.

C.5.1 Truein

C.5.2 Jibble

C.5.3 factoTime

C.5.4 AttendLab

**Reference**

[1] Mathana Gopala Krishnan, Balaji, Shyam Babu (2015), “*Implementation of Automated Attendance System using Face Recognition*”, Volume 6, Issue 3

(ISSN : 2229-5518)

[2] Mazen Ismaeel Ghareb, Dyaree Jamal Hamid, Sako Dilshad Sabr, Zhyar Farih Tofiq (2022), “*New approach for Attendance System using Face Detection and Recognition*” (Paper No. 09, 124-141)

[3] HAO YANG AND XIAOFENG HAN (2020), “*Face Recognition Attendance System Based on Real-Time Video Processing*” (Digital Object Identifier 10.1109/ACCESS.2020.3007205)

[4] Amey Shirke, Omkar Wagh, Soham Dunakhe, Vivek Shrivastav, Bhushan Mahajan (2017), “*Automated Attendance System using Machine Learning Approach”,* Volume 5, Issue 09/2017 (ISSN : 2321-0613)

[5] Qian, J., Yang, J., Xu, Y., Xie, J., Lai, Z., & Zhang, B. (2020), “*Image decomposition based matrix regression with applications to robust face recognition*” (Pattern Recognition, 102, 107204)

[6] Cuimei, L., Zhiliang, Q., Nan, J., & Jianhua, W. (2017), “*Human face detection algo- rithm via Haar cascade classifier combined with three additional classifiers*”, In 2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI) (pp. 483-487)

[7] Viola, P., & Jones, M. (2001), “*Rapid object detection using a boosted cascade of simple features*”, In Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition, CVPR 2001 (Vol. 1, pp. I-I)

[8] Zhu, Qiang & Yeh, Mei-Chen & Cheng, K.-T. Tim & Avidan, Shai. (2006), “*Fast Human De- tection Using a Cascade of Histograms of Oriented Gradients*”, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2. 1491 - 1498. 10.1109/CVPR.2006.119.

[9] Rahmad, C., Asmara, R.A., Putra, D.R.H., Dharma, I., Darmono, H. and Muhiqqin, I. (2020), “*Comparison of Viola-Jones Haar Cascade classifier and histogram of oriented gradients (HOG) for face detection*”, In IOP conference series: materials science and engineering, Volume 732, No. 1, p. 012038

[10] Geitgey, A. (2019), “*Face recognition documentation*”, Release 1.2, 3, pp.3-37

[11] Wilson, P.I. and Fernandez, J. (2006), “*Facial feature detection using Haar classifiers*”, Journal of Computing Sciences in Colleges, 21(4), pp.127-133

[12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi (2016), “*You Only Look Once: Unified, Real- Time Object Detection*”, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, doi: 10.1109/cvpr.2016.91

[13] Jiang, P., Ergu, D., Liu, F., Cai, Y. and Ma, B. (2022), “*A Review of Yolo algorithm developments*”, Procedia Computer Science, 199, pp.1066-1073

[14]