Ai approach for road Safety

University of batna 2

Allaoui Mohamed Khalil

Aouina Khansa

mars 8

Abstract

Abstract goes here

Dedication

To mum and dad

Acknowledgements

I want to thank...

Contents

List of Figures

List of Algorithms

Nomenclature

ACC Accuracy

ADAS Advanced Driving assistance Systems

BSM Blind-Spot Monitoring

DAA Driver Attention Alert

DAS Driving assistance Systems

FN False negative

FP False positive

LAS lane keep assist system

LIDAR Light Detection and Ranging

NHTSA National Highway Traffic safety administration

PR Precision rate

RADAR Radio Detection and Ranging

ROI Reagion of Interest

RR Recall rate

TN True negative

TP True positive

Chapter 1

Introduction

Recently, the development of advanced driver assistance systems (ADAS) has facilitated people's daily life from comfort to safety. However, these systems are complex [6], utilizing vehicle parameters, environmental observations, and traffic patterns to assist the driver. These systems are added cost-to-ownership due to the added expense of sensors and computing hardware needed to perceive the environment, especially in real-time monitoring. Thus, further development in this area is needed to improve reliability, performance, and decrease costs.
This work describes a driver assistance system based on computer-vision techniques.

1.1 Background and Problems

According to the World Health Organization (WHO) [4] around 1.3 million people die each year as a result of road traffic accidents, in addition to 50 million serious injuries. This cost most countries 3% of their gross domestic product.In 2016

The report also highlights More than 90% of road traffic deaths occur in low- and middle-income countries. Road traffic injury death rates are highest in the African region. Even within high-income countries, people from lower socioeconomic backgrounds are more likely to be involved in road traffic crashes.

Although current "*passive*"[1] and "*active*"[2] safety systems [21] can reduce the impact of traffic accidents, only a few car accidents are caused by bad weather and unsafe road infrastructure while most by human fault [1], such as: [4]
-Speeding

[1]Passive systems such as air-bags, seat belts, padded dashboards, or physical structure of a vehicle, normally help to reduce the severity or the consequences of an accident.

[2]Active systems like adaptive cruise control (ACC), automatic braking systems (ABS), or lane departure warning systems (LDWS) are designed to prevent or decrease the chance of crash occurrence.

-Driving under the influence of alcohol and other psychoactive substances
-Nonuse of motorcycle helmets, seat-belts, and child restraints
-Inadequate law enforcement of traffic laws

According to [5] the most likely causes of car accidents are: the driver may lose concentration on the road when driving, drivers falling asleep at the wheel, driver fatigue, or driver distraction, no matter the driver is experienced or not. A study in the United States by the National Highway Traffic Safety Administration (NHTSA) [2], confirms that almost 80% of all types of vehicle accidents involve driver fatigue, driver drowsiness, or driver distraction (in general, distracted driving), with the high speed, may cause the driver to have no time to realize the road status, which leads to car accidents.

These shocking statistics highlight the importance of research and development of advanced driver assistance systems (ADAS) focusing on *"Driver Monitoring"* by driver behavior analysis as well as *"Road Monitoring"* by road hazards detection.

1.2 Motivation

Various driving assistant systems have been developed in automotive engineering, the U.S. National Highway Traffic Safety Administration (NHTSA) defined six levels of automation from level 0 to level 5, which describes the relationship from no autonomous driving to fully autonomous driving in automotive engineering, see 1.1 .

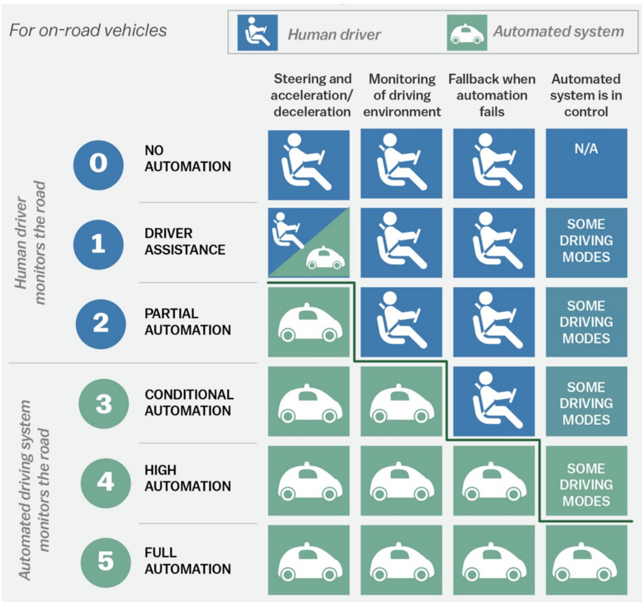


Figure 1.1: The Evolution of Automated Safety Technologies

According to the figure 1.1, in the first level, the driver needs to drive the vehicle and focus on the road to react as soon as possible. In levels 1 and 2, driving automation applies to vehicles with (ADAS) that can take over steering, acceleration, and braking in specific scenarios. But, even though level 1 driver support can control these primary driving tasks. In level 3, the system detects the environment to decide whether the driver needs to drive the vehicle, which is called conditional automation. Level 4 and level 5 indicate high automation and full automation respectively, which means the system will fully control the vehicle.

Among these levels, an (ADAS) is considered to be the basic and important component. Generally, An *ADAS* is an electronic system in a vehicle that uses advanced technologies to assist the driver [12]. It can include many active safety systems, such as [15] lane keep assist system (LAS), blind-spot monitoring (BSM), driver attention alert (DAA), and many other systems that work together to increase the safety of drivers, passengers, pedestrians, and other road users.The objective is to recognize critical driving situations by perception of the vehicle and the divers as *internal parameters*, road as *external parameters*, and the weather and lighting condition as *additional parameters.*

To collect these parameters. ADAS and autonomous driving functions feed off a continuous stream of information about the environment surrounding the vehicle, and it's the sensors' job to provide that [17] see the figure 1.2.

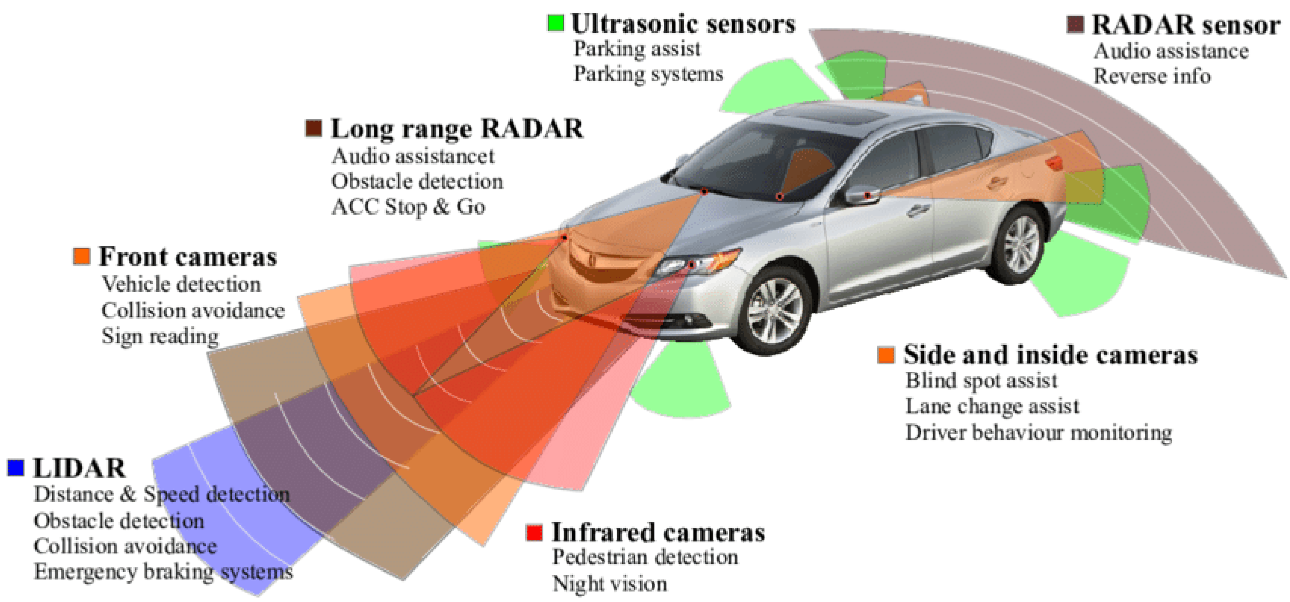


Figure 1.2: Typical types of sensors for ADAS

The three main sensors used by the automotive industry to maintain the perception for autonomous vehicles at various levels of autonomy are [17]:
Ultrasonic sensors [7], (RADAR, LIDAR) [10], Cameras [8].

Ultrasonic sensors operates by transmitting short bursts of sound waves and measuring the time taken for the sound to travel to a target object, be reflected, and return to the receiver, they are usually used for short-distance applications at low speeds, such as park assist, self-parking, and blind-spot detection.

RADAR (Radio Detection and Ranging) sensors emit radio waves and analyse the bounced wave via a receiver. Because RADAR signals can range 300 meters in front of the vehicle, they are particularly important during highway speed driving. Additionally, RADAR can see through bad weather and other visibility occlusions. Because their wavelengths are just a few millimeters long, they can detect objects of several cm or larger . LiDAR (Light Detection and Ranging) systems are used to detect objects and map their distances in real-time. Essentially, LiDAR is a type of RADAR that uses one or more lasers as the energy source. LIDARs can provide a higher resolution result but in a narrower angular field.

Camera sensors are similar to regular consumer cameras, like those that equip most smartphones. They are cheaper than both RADAR and LIDAR sensors. They can be adapted to any vehicle and any user can use them with no difficulty. For many years the fields of computer vision and image processing have used them to solve their problems. On the other hand, camera's performance drops dramatically under bad lighting conditions and they generally need a more complicated post-processing (image processing, image classification, and object detection) in order to convert the raw perceived images into a meaningful information.

Each of the above mentioned sensors have advantages and disadvantages, so that the ideal system would be a combination of all three.

1.3 Related Work

There is a wide range of research topics under the umbrella of road safety and driver assistance systems (DAS) such as *traffic signs recognition* [x], *lane detection* [x] *pedestrian detection* [x], *vehicle detection* [x], and *driver behaviour monitoring* [x] including driver fatigue, drowsiness and distraction detection. However, at a higher level, the research could be classified into two main categories: the research related to "Road monitoring" and the research works that focus on the "Driver monitoring".

1.3.1 Driver monitoring

bla bla bla

1.3.2 Road monitoring

bla bla bla

1.4 Thesis Organization

bla bla bla

Chapter 2

Theory and Concepts

In this chapter, we will simply introduce some basic concepts, methods, and mathematical background that we use in this thesis. We also provide symbols, image notations, and the equations that will be consistently used in the following chapters.

2.1 Digital Image Processing Basics

Digital Image Processing means processing digital image by means of a digital computer. We can also say that it is a use of computer algorithms, in order to get enhanced image either to extract some useful information [16].

In Digital Image Processing, signals captured from the physical world need to be translated into digital form by Digitization[1] Process. In order to become suitable for digital processing.

An image is defined as a two-dimensional function,I(x,y), where x and y are spatial coordinates, and the amplitude of I at any pair of coordinates (x,y) is called the intensity of that image at that point.An image must be digitized both spatially and in amplitude [16]. This digitization process involves two main processes *Sampling*, and *Quantization* [20].

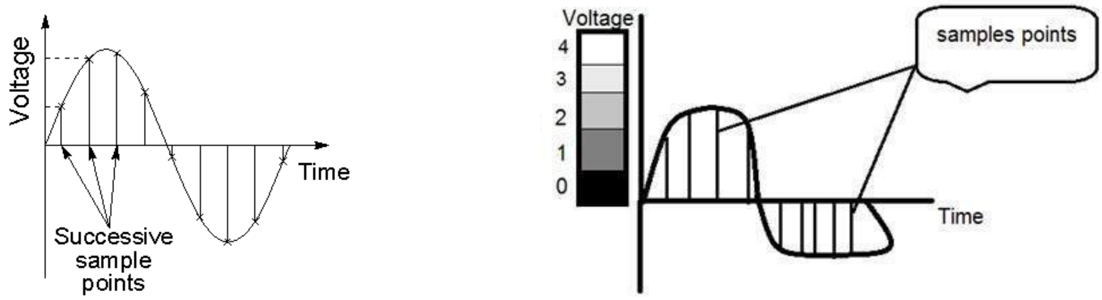
2.1.1 Sampling

In digital image processing, Sampling is the reduction of a continuous-time signal to a discrete-time signal.Since an analogue image is continuous not just in its co-ordinates (x axis), but also in its amplitude (y axis), so the part that deals with the digitizing of co-ordinates is known as sampling [18], see Figure 2.1a .

[1] Digitization: is the process of converting information into a digital (i.e. computer-readable) format.

2.1.2 Quantization

Quantization is the process of mapping input values from a large set to output values in a smaller set, often with a finite number of elements. Quantization is the opposite of sampling, It is done on the y-axis [18], see Figure 2.1b

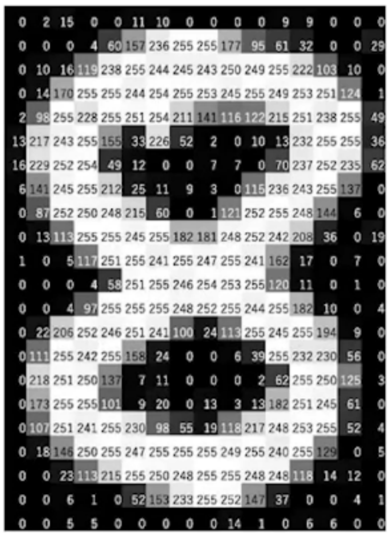


(a) Sampling

(b) Quantization

Figure 2.1: Sampling and Quantization

A digital image is typically composed of picture elements (pixels) located at the intersection of each row "**I**" and column "**J**" in each "**N**" color channels [3]. Digital images are stored in the form of a matrix of numbers where these numbers represent the intensity of each pixel, the range of these numbers (pixel values) is relative to the **Bit depth**[2], in general images are stored in 8 byte that means $2^8 = 256$ possible values.



(a) Matrix of pixels

0	2	15	0	0	11	10	0	0	0	0	9	9	0	0	0
0	0	0	4	60	157	236	255	255	177	95	61	32	0	0	29
0	10	16	119	238	255	244	245	243	250	249	255	222	103	10	0
0	14	170	255	255	244	254	255	253	245	255	249	253	251	124	1
2	98	255	228	255	251	254	211	141	116	122	215	251	238	255	49
13	217	243	255	155	33	226	52	2	0	10	13	232	255	255	36
16	229	252	254	49	12	0	0	7	7	0	70	237	252	235	62
6	141	245	255	212	25	11	9	3	0	115	236	243	255	137	0
0	87	252	250	248	215	60	0	1	121	252	255	248	144	6	0
0	13	113	255	255	245	255	182	181	248	252	242	208	36	0	19
1	0	5	117	251	255	241	255	247	255	241	162	17	0	7	0
0	0	0	4	58	251	255	246	254	253	255	120	11	0	1	0
0	0	4	97	255	255	255	248	252	255	244	255	182	10	0	4
0	22	206	252	246	251	241	100	24	113	255	245	255	194	9	0
0	111	255	242	255	158	24	0	0	6	39	255	232	230	56	0
0	218	251	250	137	7	11	0	0	0	2	62	255	250	125	3
0	173	255	255	101	9	20	0	13	3	13	182	251	245	61	0
0	107	251	241	255	230	98	55	19	118	217	248	253	255	52	4
0	18	146	250	255	247	255	255	255	249	255	240	255	129	0	5
0	0	23	113	215	255	250	248	255	255	248	248	118	14	12	0
0	0	6	1	0	52	153	233	255	252	147	37	0	0	4	1
0	0	5	5	0	0	0	0	0	14	1	0	6	6	0	0

(b) pixels values

Figure 2.2: Gray scale image of Handwritten digit

[2]The bit depth "**k**" is the number of bits per pixel, the grey scale of an image is equal to 2^k

2.2 Color Models

Generally, images that are captured by camera sensors are color images, by default they use RGB color model [19], and because in the following chapters we mainly work on grayscale images like in Figure 2.2 we need to do a color space conversion. In this section, we will introduce briefly some basic concepts related to color models used in digital image processing and color space conversion

2.2.1 RGB to Grayscale

In the RGB color model, each color in the image is obtained by superimposing three colors, i.e., red, green, blue. In this model, each pixel **P** in the image can be represented by $R(x_p, y_p)$, $G(x_p, y_p)$, $B(x_p, y_p)$. When $R(x_p, y_p) = G(x_p, y_p) = B(x_p, y_p)$ the given image becomes a grayscale image although it still has 3 color channels.

The goal is to convert the 3d [3] RGB image to a 2d[4] Grayscale image because smaller data enables developers to do more complex operations in a shorter time, there are a number of commonly used methods to convert an RGB image to a grayscale image such as average method and weighted method [14].

Average method

The Average method takes the average value of R, G, and B as the grayscale value as follows :

$$Grayscale = \frac{R + G + B}{3}$$

The average method is simple but doesn't work as well as expected. The reason is that human eyeballs react differently to RGB. Eyes are most sensitive to green light, less sensitive to red light, and the least sensitive to blue light. Therefore, the three colors should have different weights in the distribution. That brings us to the weighted method.

The Weighted Method

The weighted method, also called the luminosity method, weighs red, green, and blue according to their wavelengths. The improved formula is as follows:

$$Grayscale = 0.299R + 0.587G + 0.114B$$

[3] 3d means 3 dimensional image (width, height, depth:'number of color channels')
[4] 2d means 2 dimensional (width, height)

2.3 Image Values and Statistics

Considering **I** *a grayscale image, and X, Y are numbers of rows, columns respectively.*

Mean :

$$mean = \frac{1}{XY}\sum_{i=1}^{X}\sum_{j}^{Y} I(i,j)$$

Variance:

$$variance = \frac{1}{XY}\sum_{i=1}^{X}\sum_{j}^{Y} |I(i,j) - mean|^2$$

Energy:

$$energy = \frac{1}{XY}\sum_{i=1}^{X}\sum_{j}^{Y} |I(i,j)|^2$$

2.4 Integral Image

For a given image I, the integral image I_{int} , which was first used by Viola and Jones in computer vision [11], is the summation of all pixel values in the image, or in a window (sub-image).

Rectangle features can be computed very rapidly using an intermediate representation for the image which we call the integral image. The integral image at location (x, y) contains the sum of the pixels above and to the left of (x, y) , inclusive:

$$I_{int}(x,y) = \sum_{x'<=x,y'<=y} i(x',y')$$

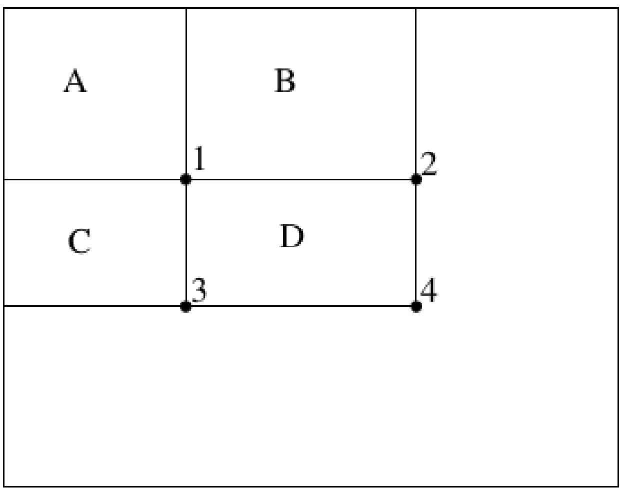


Figure 2.3: Calculation of integral values and integral image

Taking the figure 2.3 as an example, The sum of the pixels within rectangle **D** can be computed with four array references. The value of the integral image at location 1 is the sum of the pixels in rectangle **A**. The value at location 2 is **A + B**, at location 3 is **A + C**,and at location 4 is **A + B + C + D**. The sum within **D** can be computed as $4 + 1 - (2 + 3)$.

Having the integral values of each pixel calculated and saved in a data structure array, we can calculate the integral image of any image or sub-image just by applying one addition and two subtraction operations. This is very fast and cost-efficient for real-time feature-based classification algorithms, with a computational complexity of $O(N_{cols}N_{rows})$.

2.5 Classification

Classification is a process that uses a set of features or parameters to recognize an object. In this thesis we use supervised classification techniques, which means that an expert defines the classes of objects (e.g., face, eye, vehicles), and also provides a set of sample objects for a given class which is called training set. Regardless of the chosen classification technique (e.g., neural networks, decision trees, or nearest neighbour rule), we have two phases to construct a classifier: a training phase and an application phase.

Based on the provided training dataset, a classifier learns to use which set of features, parameters, and weights to be combined together in order to distinguish objects from non-objects. In the application phase, the classifier applies the already trained features and weights to an unknown query image to detect similar objects, based on what is has previously learned from the training set.

Regardless of the classification technique, the performance of a classifier can be evaluated based on the detection efficiency. Let **TP** denotes the number of true-positives or hit, when a classifiers can correctly detect the objects. Also let **FP** be the number of false-positives or miss, when a classifier wrongly detects a non-object as an object. Similarly, we can define true-negatives **(TN)** and false-negatives **(FN)** to describe correct classification of non-objects and the missing objects, respectively. Although we always can measure (count) the number of **FN**, we can not simply define the number of **TN** for an application such as vehicle detection in a road scene.

That is because the background of an image in basically uncountable. Therefore, for performance evaluation of a classifier we mainly rely on evaluations using TP, FP, and FN.

We define the precision-rate (PR), recall-rate (RR), and accuracy (ACC) as follows [13] :

$$PR = \frac{TP}{TP + FP} \qquad RR = \frac{TP}{TP + FN} \qquad ACC = \frac{TP + TN}{TP + FN + TN + FP}$$

Where PR is the ratio of true-positives to all detections by the classifier, RR is the ratio of true-positives to the total number of actual existing objects (i.e. the ground truth objects) in the analysed frames, and ACC is defined as the number of classifications a model correctly predicts divided by the total number of predictions made. It's a way of assessing the performance of a model.

2.6 Haar Feature-based Classifiers

Inspired by Haar-wavelets, Viola and Jones [11] introduced the idea of Haar-like features with square-shaped adjacent black and white patterns for the first time see Figure 2.4, in the area of face detection.

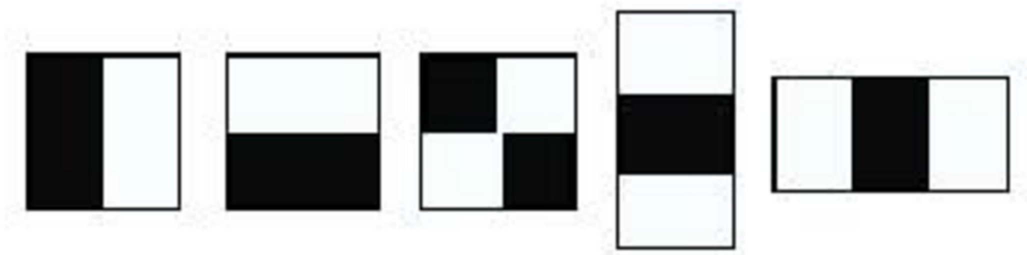


Figure 2.4: Common types of Haar-like features

The common types of Haar-like features are *Line Features*,*Edge features*, and *Diagonal features*. They are just like the convolutional kernel. Each feature is a single value obtained by subtracting the sum of pixels under the white rectangle from the sum of pixels under the black rectangle, to reduce these computations they use the integral image. These features are very important in the context of face detection because they can detect a face based on the face's properties see Figure 2.5.

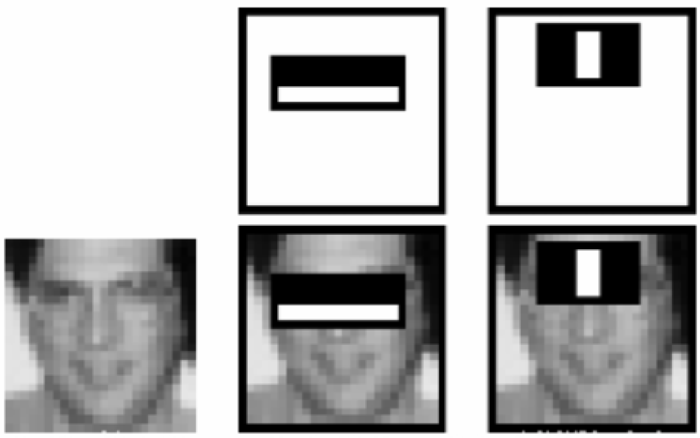


Figure 2.5: Example of Haar-feature matches in eyes

Consider the Figure 2.5. The top row shows two good features. The first feature selected seems to focus on the property that the region of the eyes is often darker than the region of the nose and cheeks. The second feature selected relies on the property that the eyes are darker than the bridge of the nose.

A *strong classifier* comprises of a series of weak classifiers (normally more than 10), and a weak classifier itself includes a set of a few Haar features (normally 2 to 5).

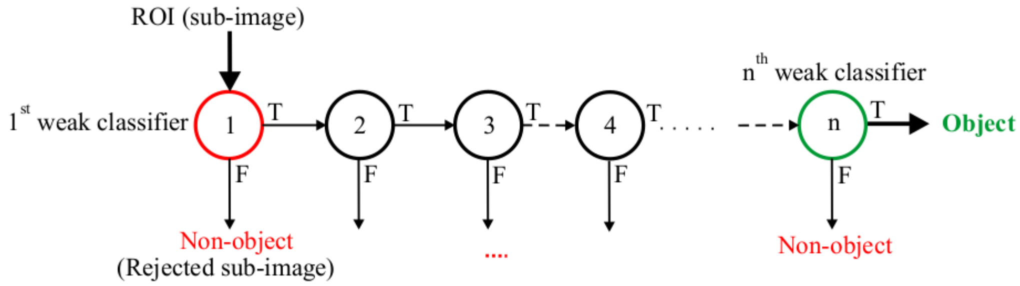


Figure 2.6: A cascade of weak-classifiers

Figure 2.6 visualizes a cascade of weak classifiers, that all together make a strong classifier. The classifier starts with the first weak classifier by evaluating the region of interest (ROI). It proceeds to the second stage (second weak classifier) if all the Haar-features in the first weak classifier match with the ROI, and so on, until reaching the final stage; otherwise the search-region under the sliding window will be rejected as non-object. Then the sliding window moves to the next neighbour ROI. If all the Haar-features within the all weak classifiers, successfully match with the ROI, then the region will be marked with a bonding box as a detected object.

The process of selecting appropriate Haar-features for each weak classifier, and then for the whole strong classifier constitutes the training phase which could be done using a machine learning technique like AdaBoost algorithm [9].

2.6.1 AdaBoost algorithm

is an approach to machine learning based on the idea of creating a highly accurate prediction rule by combining many relatively weak and inaccurate rules. The AdaBoost algorithm of Freund and Schapire was the first practical boosting algorithm, and remains one of the most widely used and studied, with applications in numerous fields [9].

Algorithm 2.1 The AdaBoost algorithm for classifier learning.

- Given example images $(x_1, y_1), ..., (x_n, y_n)$ where $y_i = 0, 1$ for negative and positive examples respectively.
- Initialize the weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for $y_i = 0, 1$ respectively, where m and l are the number of negatives and positives respectively.
- For $t = 1$ to T:

  1. Normalize the weights,
  $$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,j}}$$
  so that w_t is a probability distribution.
  2. For each feature, j, train a classifier h_j which is restricted to using a single feature. The error is evaluated with respect to
  $$w_t, \epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$$
  3. Choose the classifier, h_t,with the lowest error ϵ_t.
  4. Update the weights:
  $$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$$
  where $e_i = 0$ if example x_i is classified coreectly, $e_i = 1$ otherwise,and $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$

- The final strong classifier is:

$$h(x) = \{^{1}_{o} \quad ^{\sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t}_{otherwise}\}$$

where $\alpha_t = \log \frac{1}{\beta_t}$

Bibliography

[1] Hazrat Bilal et al. "Real-Time Lane Detection and Tracking for Advanced Driver Assistance Systems". In: July 2019, pp. 6772–6777. DOI: 10.23919/ChiCC.2019.8866334.

[2] Charlie Klauer et al. "The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data". In: (2006).

[3] Minakshi Kumar. *DIGITAL IMAGE PROCESSING*. Indian Institute of Remote Sensing, Dehra Dun.

[4] World Health Organization. *Road traffic injuries*. Fact sheet. 21 June 2021.

[5] World Health Organization. *World report on ageing and health*. Publications. 2015, 246 p.

[6] Kunal Patil. *Speeding Up the Development of ADAS Systems with Model-Based Development*. 2016. URL: https://www.dspace.com/fr/fra/home/news/engineers-insights/blog_inc_adas_1607.cfm.

[7] Paul Pickering. "Radar and Ultrasonic Sensors Strengthen ADAS Object Detection". In: (2017). URL: https://www.electronicdesign.com/markets/automotive/article/21805470/radar-and-ultrasonic-sensors-strengthen-adas-object-detection.

[8] rsipvision. *ADAS sensors: 1. RGB cameras*. 2021. URL: https://www.rsipvision.com/adas-sensors-rgb-cameras/.

[9] Robert E Schapire. "Explaining adaboost". In: *Empirical inference*. Springer, 2013, pp. 37–52.

[10] Grant Maloy Smith. *Types of ADAS Sensors in Use Today*. 2021. URL: https://dewesoft.com/daq/types-of-adas-sensors.

[11] P. Viola and M. Jones. "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*. Ed. by 10.1109/CVPR.2001.990517. 2001.

[12] Aptiv Website. *Mobility inside*. 2020. URL: https://www.aptiv.com/en/insights/article/what-is-adas.

[13] c3.ai website. *machine learning*. 2022. URL: https://c3.ai/glossary/machine-learning/classification/.

[14] Dynamsoft website. *Image Processing 101 Chapter 1.3: Color Space Conversion*. 2019. URL: https://www.dynamsoft.com/blog/insights/image-processing/image-processing-101-color-space-conversion/.

[15] Mazda Website. *Active safety technology helping prevent accidents*. URL: https://www.mazda.com/en/innovation/technology/safety/active_safety.

[16] Nishant Kumar. GeeksforGeeks website. *Digital Image Processing Basics*. 2021. URL: https://www.geeksforgeeks.org/digital-image-processing-basics/.

[17] OXTS Website. *ADAS sensors*. 2020. URL: https://www.oxts.com/adas-sensors/.

[18] Vivadifferences website. *Difference Between Sampling And Quantization In Digital Image Processing*. 2020. URL: https://vivadifferences.com/difference-between-sampling-and-quantization-in-digital-image-processing/.

[19] wikipedia website. *RGB color model*. last edit 2022. URL: https://en.wikipedia.org/wiki/RGB_color_model.

[20] Wimarshika Thamali Meduim website. *Sampling and Quantization in Digital Image Processing*. 2020. URL: https://wimarshikathamali1995.medium.com/sampling-quantization-in-digital-image-processing-8c4490357039.

[21] Wikipedia. *Automotive safety*. 2013. URL: https://en.wikipedia.org/wiki/Automotive_safety.