

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

# Automatic Nutritional Information Extraction from Photographic Images of Labels

Lara Rafaela Almeida Marinha



Mestrado Integrado em Engenharia Informática e Computação

Supervisor: António Miguel Pontes Pimenta Monteiro (Ph.D.)

Second Supervisor: Maria João Medeiros de Vasconcelos (Ph.D.)

July 21, 2015



# **Automatic Nutritional Information Extraction from Photographic Images of Labels**

**Lara Rafaela Almeida Marinha**

Mestrado Integrado em Engenharia Informática e Computação

Approved in oral examination by the committee:

Chair: José Magalhães Cruz

External Examiner: José Maria Fernandes

Supervisor: António Miguel Pontes Pimenta Monteiro

July 21, 2015



# Abstract

In the past years people showed an increasing interest in improving their diet. Many factors can be pointed to this growth, being one of them the alarming explosion of diet related diseases. This group of diseases is progressively becoming the most common causes of death, including cardiovascular diseases, obesity, diabetes and cancer.

Currently, almost all food products on the market contain nutrition labels, which is any information that appears on the product package referring to the values of the following nutrients: energy, proteins, carbohydrates, fats, dietary fiber, sodium, vitamins and minerals. This information provides a great insight of a product composition and helps the consumers to make healthier food choices.

While the labels do not have a regulated or standard format, each product often presents the nutrition information differently, leading to a wide variety of nutrition labels present in the market. This, combined with the high amount of information displayed and the difficulty of interpreting the data without the necessary knowledge, makes the extraction of relevant data and analysis a hard task for consumers. One of the solutions to simplify this task, suggested in many of the studies on this subject, is to present a summary of nutrition information as a complement to the nutrient-specific information.

The main outcome of this project is to overcome this problem and offer the consumer a tool to help in the extraction and interpretation of these values, by means of an Android application. This application tries to extract automatically the nutritional information of an image of a nutrition declaration and presents it in a single, cross-sectional shape, following the new regulations and with some additional aids, including relative values to the recommended daily doses and simplified schemes. In addition to this feature, it is also possible to compare between two products of the same category. In order to achieve these goals, it is necessary to convert the image into digital text to be processed later. To perform this conversion the application uses the OCR engine developed by Google, Tesseract.

Many problems were found throughout the development of this project, such as the low accuracy of the OCR engine or the problems of acquiring the images using a mobile device. However, after some pre and post processing algorithms, the accuracy increased to 55%, 83% more than without any preprocessing. In addition, the percentage of images that returns 0 matches decreased from 30% to 8%.



# Resumo

Nos últimos anos o interesse das pessoas em melhorar a sua dieta tem crescido. Muitos fatores podem ser apontados para este crescimento, sendo um destes o aumento alarmante das doenças relacionados com a alimentação. Este conjunto de doenças tornou-se progressivamente as causas mais comuns de morte e incluem doenças cardiovasculares, obesidade, diabetes e cancro.

Atualmente quase todos os produtos alimentares presentes no mercado incluem rótulos nutricionais, estes rótulos constituem qualquer informação presente no pacote de um produto que se refira aos valores dos seguintes nutrientes: energia, proteínas, hidratos de carbono, gorduras, fibra, sódio, vitaminas e minerais. Esta informação fornece uma visão mais detalhada da composição de um produto e ajuda os consumidores a fazer escolhas mais saudáveis.

Enquanto os rótulos não tiverem um formato regulamentado, cada produto pode apresentar a informação nutricional de maneira diferente, levando a uma enorme variedade de rótulos nutricionais presentes no mercado. Isto combinado com a grande quantidade de informação apresentada e a dificuldade em interpretar os dados nutricionais sem o conhecimento necessário, faz com que a análise dos rótulos nutricionais não seja uma tarefa fácil para os consumidores. Uma das soluções para simplificar esta tarefa, apontada em grande parte dos estudos realizados sobre este tópico, é a apresentação de um sumário da informação nutricional como complemento à apresentação exclusiva dos nutrientes e respetivos valores.

O principal objetivo deste projeto foi ir ao encontro desta necessidade e oferecer uma ferramenta que ajude na extração e interpretação destes valores, usando uma aplicação Android. Esta aplicação tenta extrair automaticamente a informação nutricional e apresenta-a de uma forma transversal, seguindo as novas regulamentações e com algumas ajudas de interpretação como valores relativos às doses diárias recomendados e esquemas simplificados. Há também a possibilidade de comparar dois produtos da mesma categoria, ajudando o utilizador a avaliar qual dos produtos é a opção mais saudável. Para atingir estes objetivos, é necessário realizar a conversão da imagem em texto para mais tarde ser processado. Para isso foi usado o motor de OCR desenvolvido pela Google, Tesseract.

Muitos problemas foram encontrados ao longo do desenvolvimento do projeto, tais como a baixa precisão do OCR usado ou os problemas derivados da aquisição das imagens usando um dispositivo móvel. No entanto, depois de alguns algoritmos de pré e pós processamento, a precisão aumentou para 55%, 83% a mais do que sem qualquer pré-processamento. Além disso, a percentagem de imagens que retorna 0 correspondências diminuiu de 30% para 8%.



# Acknowledgements

I want to thank everyone who contributed to the success of this project. To my parents, grandparents and sisters who are my biggest source of happiness and strength. Many thanks to my boyfriend and my friends, for all the help provided in this important stage of my life.

I would also like to thank to my supervisors, Miguel Monteiro and Maria Vasconcelos. Finally, I also want to leave a word of appreciation to Elsa Vieira, for her help in the nutritional aspects of this project.

Lara Marinha



*“It always seems impossible until it’s done.”*

Nelson Mandela



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Context	1
1.2	The problem and its characterization	2
1.3	Goals	2
1.4	Expected benefits	3
1.5	Structure and organization of the document	3
<b>2</b>	<b>State of Art Review</b>	<b>5</b>
2.1	Nutrition Information	5
2.1.1	Diet-related diseases	7
2.2	Nutrition Labels	9
2.2.1	Definition and types of nutrition labels	9
2.2.2	Consumer response to nutrition labels	13
2.2.3	Legislation in Portugal	15
2.3	Other applications existing in the same area	16
2.4	Mobile Development	18
2.4.1	Operating Systems	18
2.4.2	Preprocessing algorithms	20
2.4.3	Optical Recognition Character and Software	22
<b>3</b>	<b>Technical Requirements and Architecture</b>	<b>25</b>
3.1	Formal presentation of the objectives, features, use cases and requirements of the project	25
3.2	Architecture and Technologies	28
3.2.1	Packages Diagram	29
<b>4</b>	<b>Processing implementation and operation</b>	<b>33</b>
4.1	Image Processing algorithms	33
4.1.1	Extract vertical and horizontal lines	34
4.1.2	Nutrition declaration extraction	35
4.1.3	Warp Image	36
4.1.4	Divide the Image horizontally	37
4.1.5	Text location	38
4.1.6	Divide the Image vertically	39
4.1.7	Enhance the Image	40
4.1.8	Correct orientation	41
4.1.9	Post-processing	42
4.2	Final algorithm	42

## CONTENTS

4.3	Analysis of the Nutritional Information . . . . .	43
4.4	Mobile Application interfaces and features . . . . .	44
<b>5</b>	<b>Results</b>	<b>49</b>
<b>6</b>	<b>Discussion, conclusions and future work</b>	<b>57</b>
6.1	Discussion of the results . . . . .	57
6.2	Conclusions . . . . .	58
6.3	Future work . . . . .	58
	<b>References</b>	<b>61</b>
<b>A</b>	<b>Appendix 1</b>	<b>67</b>
A.1	Set of Images used in the tests . . . . .	67

# List of Figures

2.1	Example of a BOP label. Image taken from [FDA15]. . . . .	10
2.2	Example of a FoP label. Image taken from [Onl10]. . . . .	11
2.3	Example of a GDA label. Image taken from [FF15]. . . . .	11
2.4	GDA values for average women, men and children. Image taken from [FF15]. . .	12
2.5	Combination of TL and GDA label. Image taken from [HW12]. . . . .	12
2.6	Criteria for 100g of food. Image taken from [Hea13]. . . . .	13
2.7	Criteria for 100ml of drink. Image taken from [Hea13]. . . . .	13
2.8	Nutriscanner app. Image taken from [Nut11]. . . . .	17
2.9	Dominant Mobile OS. Image taken from [Sta15]. . . . .	18
2.10	Tesseract Architecture. Image taken from [PPP12]. . . . .	24
3.1	System Architecture . . . . .	28
3.2	Software architectural pattern . . . . .	29
3.3	Model Class Diagram . . . . .	30
4.1	Dilation . . . . .	35
4.2	Resultant image . . . . .	36
4.3	Intersecting points . . . . .	36
4.4	Mapping the Points to rectangle corners . . . . .	37
4.5	Nutritional label extracted with perspective correction . . . . .	37
4.6	Separate horizontally processing . . . . .	38
4.7	Separate horizontally after . . . . .	38
4.8	After processing . . . . .	39
4.9	Added border to text . . . . .	39
4.10	Separate vertically . . . . .	40
4.11	Separate vertically . . . . .	40
4.12	Original Image . . . . .	41
4.13	Image after processing . . . . .	41
4.14	Image after processing . . . . .	42
4.15	Part 1: Workflow to check the nutritional information of a product . . . . .	45
4.16	Part 2: Workflow to check the nutritional information of a product . . . . .	45
4.17	Edit the nutritional values . . . . .	46
4.18	Part 1: Workflow to compare two products . . . . .	46
4.19	Part 2: Workflow to compare two products . . . . .	47
5.1	Compare accuracy by image resolutions . . . . .	54
5.2	Compare accuracy by image resolutions . . . . .	55
A.1	cookies1 . . . . .	67

## LIST OF FIGURES

A.2 cookies2 . . . . .	67
A.3 cookies3 . . . . .	67
A.4 cookies4 . . . . .	67
A.5 cookies5 . . . . .	68
A.6 cookies6 . . . . .	68
A.7 cookies7 . . . . .	68
A.8 cookies8 . . . . .	68
A.9 cookies9 . . . . .	69
A.10 yoghurt1 . . . . .	69
A.11 yoghurt2 . . . . .	69
A.12 yoghurt3 . . . . .	69
A.13 yoghurt4 . . . . .	69
A.14 yoghurt5 . . . . .	69

# List of Tables

2.1	Reference values of energy and some nutrients . . . . .	16
3.1	Use cases . . . . .	27
3.2	Requirements . . . . .	27
4.1	Enhancements . . . . .	41
4.2	Healthier values for a yogurt . . . . .	43
4.3	Healthier values for a drink yogurt . . . . .	44
4.4	Healthier values for cookies . . . . .	44
5.1	Smartphones specifications . . . . .	49
5.2	Android results by smartphone . . . . .	50
5.3	Server results by smartphone . . . . .	50
5.4	Nutritional declaration extraction tests . . . . .	51
5.5	Warp Image tests . . . . .	51
5.6	Text location tests . . . . .	51
5.7	Enhance the Image tests . . . . .	52
5.8	Correct orientation tests . . . . .	52
5.9	Combining algorithms . . . . .	53
5.10	Summary of the averages calculated of the image processing operations . . . . .	53
5.11	Multiple images tests with different processing . . . . .	54

## LIST OF TABLES

# Abbreviations

BoP	Back of Package
CPU	Central Processing Unit
FLABEL	Food Labelling to Advance Better Education for Life
FoP	Front of Package
GDA	Guideline Daily Amount
GUI	Graphical User Interface
OCR	Optical Character Recognition
OS	Operating System
RI	Reference Intake
TL	Traffic Light
WHO	World Health Organization



# Chapter 1

## Introduction

This chapter is an introduction to the impact of nutrition labels in the human's diet and the current difficulty in reading these labels. In addition, a proposed solution to this problem is described, as well as the expected benefits.

### 1.1 Context

Food is essential for the human life, since it is the main source of energy and nutrients necessary for many vital functions. It is also an important source of health, well-being, pleasure and happiness [Bir98, Ras55]

A healthy diet not only provides the necessary nutrients to the organism to ensure a good performance, but also protects against malnutrition in all forms. However, the human's diet is changing in the last decades, becoming high in saturated and *trans* fats and salt, and low in vegetables and fruits. The previous facts combined with the tendency to a more sedentary lifestyle are considered one of the causes of the alarming explosion of diet-related diseases, such as obesity, diabetes, cardiovascular diseases and cancer [BW04, CSV<sup>+</sup>15]. These diseases progressively became the most common cause of death worldwide [WHO09]. Consequently, society has started to pay more attention to what they eat and its impact on physical and mental health [CSSM11].

Although the one's diet is an individual activity, the affordability and availability of healthier foods require the involvement of multiple sectors. For one side, the government is responsible for creating an environment that enables people to adopt healthier choices, and for another side food companies are responsible for distributing the majority of the food to the market.

In general, food associations make efforts to give some recommendations to prevent and treat these diseases, including in the majority of the cases, a healthy diet and regular physical activity. Furthermore, the World Health Organization (WHO) included the adoption of the use of nutrition labels on food products as part of its global strategy on diet, physical activity and health [WHO15].

## Introduction

As a result, almost all food products on the market contain nutrition labels. These labels usually refer to the energy values or the quantity of nutrients present, such as proteins, carbohydrates, fats, dietary fibre, sodium, vitamins and minerals. This information provides a great insight of a product composition and helps the consumers make healthier food choices.

The *Associação Fraunhofer Portugal Research* (Fraunhofer Portugal) is a non-profitable private association operating as a subsidiary of Fraunhofer-Gesellschaft, the Europe's largest application oriented research organization.

Fraunhofer Portugal is focused on the creation of scientific knowledge capable of generating valuable applications to its clients and partners, exploring technology innovations towards economic growth, the social well-being and the improvement of the quality of life of its end-users.

This project was developed to be integrated in the bigger Smart Companion project. Many projects are developed to meet the older adult's needs being Smart Companion one of these projects. This project offers an Android customization in order to simplify the use of mobile devices for inexperienced users and be a companion to support the users in their daily actions: preventing isolation, promoting autonomy and quality of life, and improving health management.

### **1.2 The problem and its characterization**

Nutrition labels offer the advantage at the moment of purchase of food to provide information about the composition and nutritional information. However, although nutrition labels are a very important tool, studies suggest that the use of this information is actually very low, due to difficulties in reading and interpreting them.

In Portugal, the presence of nutrition labels on products only became regulated at the beginning of this year, but the manufacturers have an extended limit to comply until 2016. Although there are mandatory terms, units of measurements and minimum line height established, a standard format does not exist.

The variety of formats of nutrition labels that can be found combined with the high amount of information presented makes the reading of these labels very confusing to the consumers. So, to the common consumer, it is hard to understand and consequently extract relevant data.

Besides that, it is not enough to look at the food labels, it is also important to know about the nutrients and its impact in the overall diet specific knowledge, in order to interpret the nutritional data presented and indeed make healthier choices.

### **1.3 Goals**

In this work it is intended to develop a mobile application that helps the user to interpret the data presented in the nutrition labels.

After the user acquires the image of the label with nutritional information, the application should analyze it using image processing algorithms and an OCR engine to automatically extract

the nutritional information. The goal is to display the relevant information on a single, cross-sectional shape, so the information presented is relevant, accurate and easy to understand.

The application should also help the user to compare between two identical food products, by choosing the best option and assessing the suitability of a specific product according to personal information.

### **1.4 Expected benefits**

This application aims to help the users to improve their health, by affecting their food choices to healthier options. Also it should allow the user to assess the quality of a product, its benefits and possible risks concluding if it is appropriate for his specific diet.

It aims to help users better understand the nutritional information available in food labels having a better comprehension how each food fits in the user's overall diet.

### **1.5 Structure and organization of the document**

This document is structured in six chapters. This chapter presents an overview about the context of this project, the characterization of the problem that needs to be solved and its solution. It also presents an overview about the expected goals and benefits of that solution.

The chapter 2 introduces the areas where this project will be developed. Initially, it is described in detail the nutrition labels and its roles in the people's lives. Next, an introduction to some image processing operations and some possibilities of text extraction. In addition, some products already developed in the same area and that compete with this work are also presented.

In the chapter 3 is formally presented the solution and its objectives. The technical requirements and architecture of this project are also described.

The detailing of implementation the solution is presented in the chapter 4. First is described the technical part of this project that includes the extraction of the text and the image processing operations performed, and after the analysis of the nutritional data collected.

In the chapter 5 is presented the tests performed to evaluate the image processing operations. An analysis of the results is also presented.

Finally, the chapter 6 presents a summary of the analysis of the tests, and concludes this project presenting some possibilities for future work in order to solve some issues found during the development of this project.

## Introduction

## Chapter 2

# State of Art Review

This chapter includes a detailed characterization of the nutrition labels and its impact in the overall diet as well as introduces some scientific themes and trends that make possible the extraction of text of an image using mobile devices. Some products already developed in the same area and that compete with this work are also presented.

### 2.1 Nutrition Information

Alimentation is essential in human's life being the main source of energy and nutrients necessary for vital functions like growing, moving, reproducing or maintaining of body temperature [Bir98, Ras55].

A balanced diet can provide the energy and nutrient requirements to the organism while ensuring a good performance and preventing several diseases [CSV<sup>+</sup>15, Wil99]. In contrast, a poor diet may have a negative impact on human health.

Besides being an important source of health, it is also an important source of happiness. In fact, several studies show that specific nutrients can stimulate certain glands or hormones responsible for the happiness feeling, and also have a deep influence in the overall treatment of depression [AGH06, Ber96]. Furthermore, eating and drinking are considered two of the major occupancies of waking time, consuming in average 1 hour per day, without including preparing meals, shopping and cooking. Although these are activities that can be done alone, meals are usually shared among family, friends or co-workers, being as well an important part of social life.

According to the data collected from the World Health Organization (WHO), between 2000 and 2012 the leading causes of worldwide mortality are diet-related chronic diseases, with the number of deaths growing every year [WHO09]. In Portugal, according to the results from 2004-2010 Portugal National Health Plan, the three main causes of death are also linked to the diet-related diseases: cancer, cardiovascular diseases and diabetes [Sau13].

## State of Art Review

Consequently, in the past years people show an increasingly interest in improving their diet [CSSM11, MHVCV10, CS05]. Eating healthy became a really important task in the daily life. However, making healthy and informed decisions is not easy, since it means choosing many different foods throughout the entire day to get all nutrients needed [GW07].

Nutrition science offers some guides to educate people about eating healthier and gives some advice, on matters of the interaction of nutrients and other substances in food. There are five key terms that often appear in nutrition facts and are important to understand the role of each parameter to the overall diet: Calorie, Fat, Sugar, Salt and Protein.

A calorie is an unit of measure of energy obtained by macronutrients, which includes carbohydrates, fats, proteins and alcohol. One calorie is formally defined as the amount of energy necessary to increase by 1°C the mass of 1 gram of water, and corresponds to approximately 4.2 joules. Calories are usually used in nutrition contexts to measure the quantity of energy obtained from a food substance or to help manage weight, They are commonly represented in kilojoule (symbol: kJ), kilocalories (symbol: kcal) or Calorie (symbol: Cal).

Fat is a high source of energy and each gram of fat contains nine calories. It also adds texture, flavour and aroma to food. It is an important macronutrient with structural and metabolic functions: improves insulin signalling, increases intestinal absorption of lipophilic vitamins and maintains a healthy skin and hair. As fat is a high supplier of energy in a compact form, it should be consumed moderately, not more than a third portion of a diet. Chemically, fat molecule is composed of glycerol and different types of fatty acids. Depending on the number of double bonds in the molecular structure of fatty acids, fats can be classified as saturated, monounsaturated or polyunsaturated. Saturated fats have fatty acids without double bonds in its structure and can be found in whole milk products, red meats and some vegetable oils (palm oil and coconut oil). The consumption of saturated fats needs to be limited because a high intake can raise LDL-cholesterol levels, leading to dyslipidaemia and other cardiovascular disease, as well to obesity. Unsaturated fat is composed of fatty acids that have one (monounsaturated fats) or more double bonds (polyunsaturated fats) in its structure. This type of fat is healthier than saturated fats because they do not raise LDL-cholesterol and can even increase HDL-cholesterol, preventing the development of cardiovascular diseases. Unsaturated fats can be found in nuts, vegetable oils (olive, peanut and canola oils), fish and white meat. Although they have a natural source and present some benefits, are still high in calories and should be consumed with moderation. Another type of unsaturated fats is the *trans* fat. Although they appear in natural products, *trans* fats are mainly created through an industrial process that adds hydrogen to vegetable oil. The consumption of this type of fat is highly discouraged because it can raise the levels of LDL-cholesterol and lead to many diseases.

Sugar is also called as a “carbohydrate simple”, it means carbohydrates that are rapidly absorbed by the organism, contributing to normal glycemic establishment. Like “complex carbohy-

drates” (starch, cellulose, chitin), each gram of sugar simple carbohydrates contains four calories. Carbohydrates are essential to the brain function, a high source of energy and prevent the use of protein to energy so it can be used to build and repair the body. Natural sugars, as fructose, lactose, sucrose and glucose, can be presented in several food sources, as fruits vegetables or milk products. To a healthy diet, the intake of sugars should be limited, otherwise can lead to obesity, diabetes and tooth decay.

Salt is a mineral substance composed of sodium and chloride. It is an important nutrient because helps the regulation of the body’s water content and is involved in the nerve function, muscular activity and energy usage. Consumed in excess can increase the blood pressure, leading to hypertension problems. Thus, the intake of this mineral must be limited.

Proteins are macromolecules composed of amino acids. They are responsible for the transport of nutrients and oxygen to cells, making these macronutrients essential to all the cells of the body. Proteins also supply energy where each gram contains four calories, and are responsible for the growth and repair of body tissues. Some amino acids, called as non-essential amino acids, can be synthesized by the human organism. However, some amino acids cannot be synthesized by the organism (called as essential amino acids), so they must be obtained from proteins present in animal and plant foods. The intake of proteins is very important in our diet.

### **2.1.1 Diet-related diseases**

The World Health Organization recommends a healthy diet including the consumption of fruits, vegetables, legumes, nuts and whole grains for an average adult. It is recommended that the daily diet contains less than 10% of total energy from free sugars and less than 30% of total energy from fat. The intake of unsaturated fats is preferable to saturated fats and the consumption of industrial *trans* fats is highly discouraged.

In the past years, human’s diet suffered a massive change because of the change in food availability, food prices and level of income. Traditional plant-based diets are being constantly substituted by diets rich in added sugars and animal fats and low in dietary fibre, fruits and vegetables [CSV<sup>+</sup>15]. This, combined with a general trend towards a more sedentary lifestyle and low level of physical activity, increases the risk of developing chronic diseases [CSV<sup>+</sup>15].

Diseases that are diet-related are growing and have high mortality rates, being on the top of the causes of death worldwide [HEL<sup>+</sup>04, NKP99]. Some of these diseases are cardiovascular diseases, obesity and diabetes, among others.

Cardiovascular diseases (CVDs) are a group of disorders of the heart and blood vessels that can lead to death. It is the number one cause of death globally and includes numerous problems,

many of them related to atherosclerosis, in which plaque (made of fat, cholesterol, calcium and other substances) builds up inside of the arteries putting the oxygen-rich blood flow in risk. If the blood flow stops, it can cause a heart attack or a stroke. To prevent CVDs it is recommended to have a regular physical activity, avoid tobacco, have a diet rich in fruit and vegetables and avoid foods that are high in fat, sugar and salt. CVDs are among the principal causes of morbidity and mortality in Portugal, particularly stroke and ischemic heart disease, which are the leading cause of death in both sexes [RFP13].

Overweight and obesity are defined as abnormal or excessive fat accumulation that presents a risk of death, respectively. A risk factor for obesity is the total calorie intake (especially from fats and sugars) higher than the expended daily and sedentary lifestyles. Body Mass Index (BMI), is a widely parameter used to measure the risk of obesity; is calculated dividing the weight (in kilograms) by the square of the height (in meters). An adult who has a BMI of 30 or higher is classified as obese.

Obesity is a chronic disease with a high risk of death and is one of the principal causes of death worldwide, continuing to spread all over the world, even in low and middle income countries [WHO00]. It is classified as a complex disorder because can also lead to other diseases like cardiovascular diseases, diabetes, musculoskeletal disorders and cancer (endometrial, breast and colon) [WHO00]. Obesity treatment includes dietary changes and increased physical activity. Dietary Guidelines and recommendations for obesity prevention limit the total fats and sugars and the consumption increase of fruit and vegetables.

Diabetes is a chronic disease that occurs when the pancreas does not produce enough insulin (hormone that regulates the blood sugar) or when the body cannot effectively use the insulin it produces. In this last case, called type II diabetes, dietary factors (and obesity) are the main causes of this disease development. Type II diabetes is the most common type of diabetes, corresponding up to 90% of cases. It's a worldwide problem that can lead to death, but occurs essentially in low and middle income countries. In Portugal, for the next 20 years, the prevalence of diabetes was estimated at 11.7%, being more frequent in men (14.2%) than women (9.5%) and increasing with age [CCI13].

Diabetes can be prevented by adopting a healthy diet and increase physical activity. It is highly recommended the reduction of quantity of *trans* and saturated fats and the increase of the quantity of healthier unsaturated fats, but even this type of fats should be consumed with moderation. The National Diabetes Information Clearinghouse advises to control the portions consumed where a small or medium-sized man should ingest 1200 to 1600 calories daily, and medium-sized or large man should ingest 1600 to 2000 calories, including eating 3 to 5 servings of fruit and vegetables a day. Also recommends avoiding high-sugar, high-calorie and high-fat desserts. Many of these products have a "light" alternative, with low calories and low percentage of sugar.

The increase of diabetes motivated the research and development of "light" products, but these products are susceptible to loss of colour, flavour and texture [GZMS05]. To overcome this prob-

lem companies often replace sugars by sweeteners. Sweeteners are food additives that substitute sugar, adding the same taste usually with less food energy. Some sweeteners can be found in nature while others are synthetic. If used correctly by the manufacturers, can help type II diabetics controlling their weight and decreasing the consumption of sugar [RMD<sup>+</sup>09], otherwise can worsen health consumers.

## 2.2 Nutrition Labels

To help the consumers making healthier choices at the moment of purchase, in 1991, the Committee on Medical Aspects of Food Policy, set reference values for the intake of energy, protein, fats, sugars, starches, non-polysaccharides (NPS), vitamins and minerals. Seven years later, a set of guidelines for labelling purposes was developed as a way to inform the consumers of the recommendation of the government about nutrient values.

Nowadays, most of the countries do not have the nutrition labels regulated, however a high percentage of products in the market have nutrition labels [gBCL<sup>+</sup>10] making them one of the primary sources of nutritional information. By definition, the nutrition label is regarded as any information displayed on the product package. Usually, product packages have some references to energy values or to the following nutrients: proteins, carbohydrates, fats, dietary fibres, salt/sodium, vitamins and minerals.

Even with new regulations only emerging now in Europe, the consumers' interest in a balanced diet and differentiation between products forced the manufacturers to adopt some form of nutrition labels [CSSM11].

### 2.2.1 Definition and types of nutrition labels

The nutrition label works as an identity card of a food product and its main purpose should be provide to the consumer useful and trustful information at the point of purchase.

Consumers can use these labels for different reasons: checking the nutritional content of a product, compare similar food products or choosing the best food product to the specific needs and diet restrictions. Checking the nutrition labels regularly makes the consumers more conscious of what they are eating, the role of the nutrients in their overall diet and how to prevent some diseases [Lee03]. Thus, these labels are a valuable tool for the customer to help maintaining a healthy diet and encourage the consumption of healthy products.

Usually the nutrition labels include the ingredient list, nutrition declaration and optional health or nutritional claims. The ingredient list is composed by each ingredient of a food product in descending order by weight. Health claims describe the relation between a food substance and a specific nutrient, the reduced risk of a disease or health-related condition. Some examples are: "free of sugar", "gluten free" and "prevents cancer", among others. Health claims should follow a set of global nutrition standards and because it affects the nutrition judgment about a food product, the validity it's very important. The information about the energy and nutrients can be presented in

several formats in the food market: Guideline Daily Amount (GDA) labels, Traffic lights, health logo, specific claim, nutrition grid, ingredient list, color-coded GDAs, among others.

The nutrient values can appear per grams (g), millilitre (ml) or serving unit. A serving unit is the amount of food that is generally served and for which the nutritional values are referred. These values are based on the average amount of what people eat of a certain product, but it's not an accurate value. Depending on the type of product, these units can be presented in different formats, such as cup, tablespoon or number of units. Also, the recommended daily values of each nutrient are presented using GDA or Reference Intakes (RI). The nutrition declaration can be divided in two main categories: Back of Pack (BoP) or Front of Pack (FoP).

BoP consists in detailed nutritional information presented in the back of the package of a product. Usually it is displayed in a tabular form, as a panel or grid, with the numbers aligned or presented linearly. All the nutritional information, including energy and each nutrient, is presented per 100g or 100ml and/or per serving unit. An example of BoP label is presented in Figure 2.1.

It is also usual to find the percentage of the recommended daily amount of calories, fat, saturated fat, total sugars and sodium/salt. According to Food Labelling to Advance Better Education for Life (FLABEL), an EU-funded research project, BoP tabular or linear was the most popular format. Results showed that 85% of the products audited in all the EU member States and Turkey present this format of labels [gBCL<sup>+</sup>10].

<b>Nutrition Facts</b>	
Serving Size 2/3 cup (55g)	
Servings Per Container About 8	
Amount Per Serving	
<b>Calories</b> 230	Calories from Fat 40
% Daily Value*	
<b>Total Fat</b> 8g	<b>12%</b>
Saturated Fat 1g	<b>5%</b>
Trans Fat 0g	
<b>Cholesterol</b> 0mg	<b>0%</b>
<b>Sodium</b> 160mg	<b>7%</b>
<b>Total Carbohydrate</b> 37g	<b>12%</b>
Dietary Fiber 4g	<b>16%</b>
Sugars 1g	
<b>Protein</b> 3g	

Figure 2.1: Example of a BOP label. Image taken from [FDA15].

FoP summarizes the nutritional information in the front of the package, using symbols and rating systems to present a summary of the key nutritional characteristics of a product [HWA<sup>+</sup>13]. This type of labelling can be very useful to compare different food products quickly. Usually FoP

presents the nutrition values of energy in kJ or kcal and the amount of fats, sugar and salt in g, considering a serving unit of food. They can have a lot of formats, some examples of FoP labels are presented in Figure 2.2.

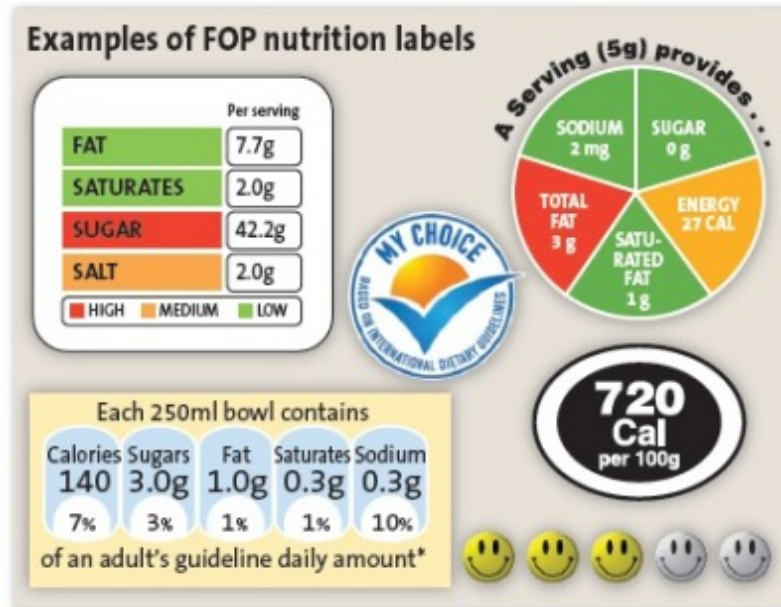


Figure 2.2: Example of a FoP label. Image taken from [On110].

It is not enough to provide food to the organism, being the amount in which it is consumed very important. The UK government in collaboration with the food industry and consumer organizations provided guidelines to help the consumers understanding nutritional information presented on the food labels. GDA labels present these guidelines about the approximate amount of calories, fat, saturated fat, total sugars, and sodium/salt, a healthy adult or children should ingest per day. One example of this type of label is presented in Figure 2.3.

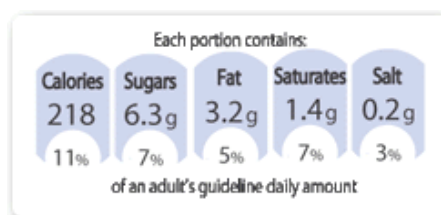


Figure 2.3: Example of a GDA label. Image taken from [FF15].

GDA values are not regulated, but despite that, values are broadly consistent between UK, Europe and internationally [FDF14]. One example given by the Food and Drink Federation in UK is presented in the Figure 2.4.

Recently, this term has been replaced by “Reference Intake” (RI). Like GDA, RIs are recommendations of the quantity of energy and nutrients that should be daily consumed. This is the first

time that the actual values are set and regulated by European laws. Unlike GDA, only values about the mandatory ingredients and energy are defined and only one set of RIs is defined for average adults instead of values for men, women and children.

Guideline Daily Amount Values			
Typical values	Women	Men	Children (5-10 years)
Calories	2,000 kcal	2,500 kcal	1,800 kcal
Protein	45 g	55 g	24 g
Carbohydrate	230 g	300g	220 g
Sugars	90 g	120 g	85 g
Fat	70 g	95 g	70 g
Saturates	20 g	30 g	20 g
Fibre	24 g	24 g	15 g
Salt	6 g	6 g	4 g

Figure 2.4: GDA values for average women, men and children. Image taken from [FF15].

Traffic Lights (TL) labels are colour coded labels that display a ranking of total fat, saturated fat, sugar, salt, and sometimes energy, along with the amount of grams of each in one serving. Normally it is easier for the consumer to understand and interpret the nutrient composition using this concept [HWA+13]. This system is an UK-wide government scheme in which the levels high, medium and low correspond to the colours red, yellow and green, respectively. In short, food products with more green lights are healthier choices.

Combinations of TL and GDAs can also be used, as shown in Figure 2.5. These two formats provide an overall nutritional score unlike the nutrient-specific systems.

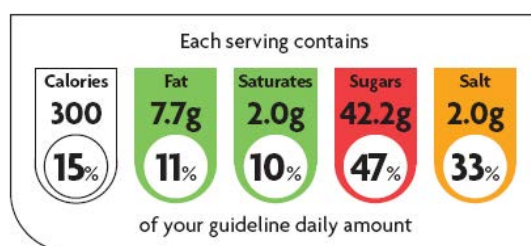


Figure 2.5: Combination of TL and GDA label. Image taken from [HW12].

There aren't regulated values for this type of labels in Portugal. Although the Department of Health offers some criteria combining the colour coding and the requirements of Article 35 of EU Regulation 1169/2011 [Hea13]. The criteria for food is presented in the Figure 2.6 and for drinks in the Figure 2.7.

Text	LOW	MEDIUM	HIGH	
Colour code	Green	Amber	Red	
Fat	≤ 3.0g/100g	> 3.0g to ≤ 17.5g/100g	> 17.5g/100g	> 21g/portion
Saturates	≤ 1.5g/100g	> 1.5g to ≤ 5.0g/100g	> 5.0g/100g	> 6.0g/portion
(Total) Sugars	≤ 5.0g/100g	> 5.0g and ≤ 22.5g /100g	> 22.5g/100g	> 27g/portion
Salt	≤ 0.3g/100g	> 0.3g to ≤ 1.5g/100g	>1.5g/100g	>1.8g/portion

Figure 2.6: Criteria for 100g of food. Image taken from [Hea13].

Text	LOW	MEDIUM	HIGH	
Colour code	Green	Amber	Red	
Fat	≤ 1.5g/100ml	> 1.5g to ≤ 8.75g/100ml	> 8.75g/100ml	>10.5g/portion
Saturates	≤ 0.75g/100ml	> 0.75g to ≤ 2.5g/100ml	> 2.5g/100ml	> 3g/portion
(Total) Sugars	≤ 2.5g/100ml	> 2.5g to ≤ 11.25g/100ml	> 11.25g/100ml	> 13.5g/portion
Salt	≤ 0.3g/100ml	>0.3g to ≤0.75g/100ml	> 0.75g/100ml	> 0.9g/portion

Figure 2.7: Criteria for 100ml of drink. Image taken from [Hea13].

## 2.2.2 Consumer response to nutrition labels

Several studies about the relationship between consumers and food labels were made and analysed in the last years, mainly in USA and UK. In Europe, the impact of the food labels has been intensely studied in UK because it has been the leading promoter of food labels, but recently, studies are widely spreading in other parts of Europe. These studies show that consumers like the idea of simplified front pack information but reveal many problems in understanding product labels [GW07, CS05, GFCW<sup>+</sup>10].

Analysing and extracting information through these labels is normally a hard task for the customers. Many causes are pointed out and described through the different studies, some converging to the same result.

One cause pointed out is the wide variety of different formats resulting from the lack of a standard format. Consumers considered that interpreting and comparing labels of products with different formats was time-consuming and inconvenient. This inconsistency increases the consumers' confusion and lack of interest. FLABEL research project explains that the lack of motivation and

attention among consumers must be overcome by providing nutrition labels easier to access while shopping [gBCL<sup>+</sup>10].

Other cause presented is that the nutritional information is not always comprehensive. The use of technical language and terms is one of the reasons identified by [CS05]. This author claims that basic concepts like 'fat', 'calories/kilocalories', 'sugar', 'vitamins' and 'salt' were well understood by consumers, but some relations between the nutrients and the contribution of each nutrient in the consumer's diet was least understood. Consumers seem to be able to understand certain key terms and also apply this information in simple tasks like extracting and understanding simple information, making simple calculations and comparing nutritional characteristics of products. But show difficulty in understanding the role of the nutrients in their diet, converting "g per 100g" to "g per serving" and understanding the concept of serving size information [CS05]. Also, different results were found when the information of the nutrition label is applied for judging different products. For example, when regarding to single nutrients or the overall healthiness, most of the consumers enquiries come up with the right response [CS05, GFCW<sup>+</sup>10]. However, the best results were found when the comparison was about well-known nutrients (like 'fat') or between similar products.

A correlation between reading nutrition labels and demographic, socioeconomic and ethnic factors has also been reported in some studies. In general, these studies have shown that women use food labels more than men to make food choices [CS05, GW07, NKP99]. Also consumers who believe in the relationship of food with preventing diseases, e.g. athletes, are more likely to use food labels. Results of an ACNielsen study, conducted in 38 countries, showed a low percentage in checking nutritional information in almost every country. Portugal, in contrast, showed one of the highest rates when asked about the use of nutrition labels; 44% of the consumers enquiries claimed that "always" check the nutritional information on the package [GW07]. FLABEL research project is also in agreement with this result, in which FoP nutrition labels reached 37% of use in Portugal, being one of the leading European countries in penetration of labels FoP and BoP [gBCL<sup>+</sup>10].

The use of nutrition labels also varies across different types of situation and food. According to [CS05], when there is a time limit to shopping, consumers do not express interest in checking nutrition labels, which include most of the shopping experiences. In the other hand, consumers show a bigger interest in the nutritional values in pre-packaged food and breakfast products. The taste of food is the main interest of the consumers when choosing between products, followed by 'this is what my family wants', price/special and at last health/nutrition [GFCW<sup>+</sup>10].

In conclusion, after analysing all the reviews and studies, it is clear that a solution that makes the reading and understanding of nutrition labels easier and more practical is necessary and welcome. There is no consensus among the customers about which format (single and simple against complex and detailed) is most helpful to understand the nutritional information of a product. Most studies conclude that a summary of the nutrient characteristics is more helpful for the consumers

than nutrient-specific information. Adding interpretation aids is referred as a solution like verbal descriptors, recommended values or types of benchmarks [CS05]. Another option consists in complementing the complex and detailed information on the back of the package with a summary, resuming the principal nutrition characteristics, in the front of the package [GW07].

### 2.2.3 Legislation in Portugal

Nowadays, most of the consumed food is supplied by the food industry, making the nutrition control more difficult [Nes13]. As food producers don't work with established rules, they can choose to present whatever they want and the way they want into food labels, without having into account the consumers' interests. Thus, legislation of the nutrition labels is the response to protect the consumers by restricting the information presented. The importance of food labels legislation is to guarantee that consumers have a more complete and accurate information about the content and composition of a food product. Nutrition labels should work as a channel between the food companies and the consumers.

Until the 13th of December of 2014 the nutrition labels in Portugal did not have any legislation, but after that date the Regulation (EU) 1169/2011 of the European Parliament and of the Council of the 25th October was implemented. These rulings set a wide-changing enforcing to the layout and information contained in labels, in order to simplify and harmonize the nutritional information displayed on the food product's package, making the most valuable information available.

These rules are mandatory for all the food products that are produced, commercialized and packed in the consumer's absence and ready to be supplied, but companies have an extended deadline for compliance until the 13th of December of 2016.

Food labels must present the name of the product, the ingredient list, the nutrition declaration, the liquid content, indication of the origin, name or social reason and address of the manufacturer/importer, number of the lot, expiration date and instructions about how to prepare and use, when needed. The height of the characters should be equal or higher than 1.2 mm, except for small packages. The information should not contain any information that may mislead the consumer into error, confusion or mistake. It should be exact, clear and easily comprehended by the consumer.

The nutrition declaration is mandatory in the back of the product package with the following terms: Energy (kJ/kcal), Lipids (g) of which saturated fatty acids (g), Carbohydrates (g) of which Sugars (g), Proteins (g) and Salt (g). The amount of each nutrient/energy needs to be expressed using absolute values and listed in a certain order - energy (both in kilojoules and kilocalories) and the amounts of fat, saturates, carbohydrates, sugars, protein and salt (in place of 'sodium' which is no longer permitted). If other nutrients are included they should be in the same visual field. This information should be presented in tabular form with the number aligned, but if there is no space it can be presented linearly.

In the nutrition declaration, reference values can only be presented about the required terms, and only if it is presented for each one. Also, the information in the nutritional declaration can be voluntarily repeated in a FoP format, and only reference values about energy can be presented or, as in the nutrition declaration, all the required terms. To calculate the energy and nutrient contents present in pre-packed food products, “Reference values” for adults should be considered using the regulated values and shown on Table 2.1 [EU11].

Table 2.1: Reference values of energy and some nutrients

Energy or nutrient	Reference values
Energy	8 400 kJ/2 000 kcal
Total Lipids	70 g
Saturated fats	20 g
Carbohydrates	260 g
Sugars	90 g
Proteins	50 g
Salt	6 g

### 2.3 Other applications existing in the same area

There are many mobile applications regarding to nutrition data and its influence on the overall diet. Nutriscanner, Mobile Food Information Scanner and Automatic Grocery Shopping Assistants are some examples of mobile applications that perform optical character recognition to extract nutrition information from the food labels.

Nutriscanner is an iOS application with Optical Character Recognition (OCR) to scan nutrition Facts Labels directly from photos taken with an iOS device. One of the reasons for developing for iPhone is the high resolution cameras used in these types of devices. As the picture resolution increase, the possibility of easily extracting more reliable information also increases. Also the ARM processor together with artificial intelligence is capable of interpreting the information in seconds. The nutritional information attributes extracted include serving size, weight, calories, fat, carbohydrates and protein. An indicator is displayed to show the user if the extraction of information about each attribute was successful. If it was successful, the indicator is presented with the colour green, otherwise the colour is red and the user need to enter the data manually. A screen of Nutriscanner is shown on the Figure 2.8. Nutriscanner points out a study of Bidisha Mandal [Man10] that suggests that reading nutrition labels helps people to lose weight. Also results suggested that between people who only performed one activity, label-reading or exercising, label readers presented an advantage in weight loss.



Figure 2.8: Nutriscaner app. Image taken from [Nut11].

Mobile Food Information Scanner is a mobile user interface designed to extract semi-automatically information presented in the ingredient list of a nutrition label. It is composed by an OCR server and a mobile user interface that can be used to capture images of the labels using an Android device, send them to the server to be processed and present the retrieved information. This information can be validated after by the user, directly in the mobile device. If the results are not correct, a detailed correction view is available with alternatives to the user selects. To extract the ingredient information an open-source OCR tool is used, the Cuneiform, that also provides bounding boxes of recognized letters. Some preprocessing and post-processing techniques are performed to achieve reasonable results and overcome the problem of poor images [LSAR00].

Automatic Grocery Shopping Assistant is an Android mobile phone application that keeps track of common groceries purchased and their associated nutritional values. It is composed by an OCR server that will perform all the image processing tasks. One of the main features of this application is to add items to a custom library and also the nutrition facts associated just by taking a photograph during a purchase [YY11]. The user takes a photograph of the nutritional information on the package of a product using a mobile interface that sends that image to the server. The server will process the image and using the OCR Tesseract program will extract the information about the serving size and calories.

## 2.4 Mobile Development

By the end of 2013, according to the market research firm eMarketer, almost 20% of the world population owned a smartphone, and it is estimated that by 2017 this number will raise to one third [eMa14]. In most developed countries, this percentage is even higher. In May of 2013, according to another eMarketer study, in America almost 60% of people have a smartphone and in Europe, some countries like Norway and UK reach up to more than 60% [McD13]. Smartphones offer a large number of functionalities [LLKK05] for a low price [BGR07] such as: taking pictures, web browsing, access to internet and Wi-Fi, GPS navigation, media players, mobile payments, motion sensors, between others. Also a large variety of applications like games, scheduling, social networks, are easily available in online stores. There are many operating systems in the market: Android, iOS, Windows Phone, Firefox OS, Sailfish OS, Tizen, Ubuntu Touch, BlackBerry, among others, being the most common Android and iOS. A study performed by StatCounter in November of 2013, shows that although Android has 43% of the worldwide market share, its dominance is not consensual in all countries. iOS share an important part of this market, leading in some countries, as shown in Figure 2.9 [Mah13]. Most of the countries in which iOS is popular are developed countries and the majority of the consumers have high incomes. According to the website StatCounter, from November of 2013 until now, Android popularity increased to almost 60% [Sta15].

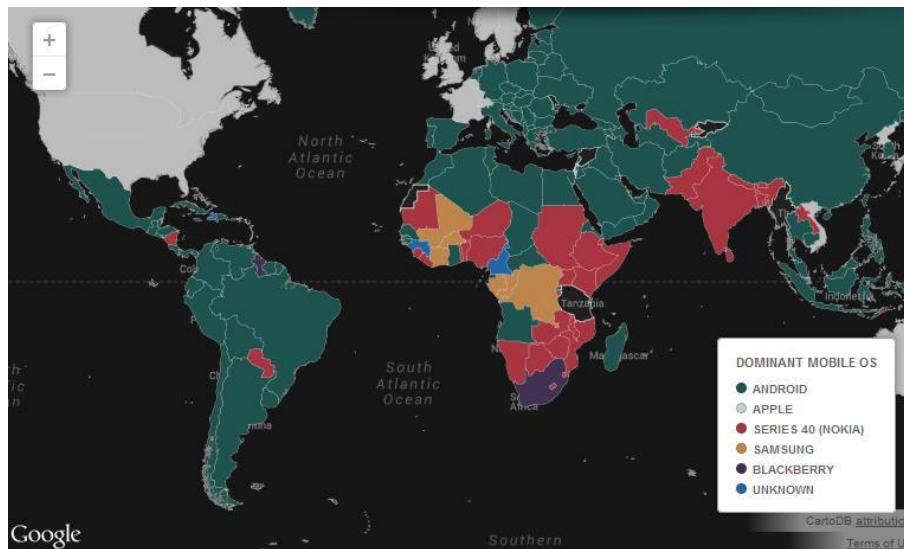


Figure 2.9: Dominant Mobile OS. Image taken from [Sta15].

### 2.4.1 Operating Systems

Android is a software platform and operating system designed mainly for smartphones and tablets. It was initially released in 2008 and it is currently developed by Google and Open Handset Alliance. This operating system is open source and based on the Linux kernel, that provides a base

operating system, an application middleware and a Java software development kit which can be found also in smartwatches, televisions, glasses and cars.

Android is the leading operating system in the market, having a large community of developers. One of the reasons for this success is the number of different devices and companies that support and promotes Android, resulting from the Open Handset Alliance partnership, which is a major consortium of 48 companies, including Samsung, Motorola, T-Mobile, Intel and Qualcomm. Another reason is the open business approach offered by Google that gives freedom and flexibility to the developers, resulting in a high number of different applications available to the users. But this also raised some problems, like the need of developers to have into account all the different requirements and or limitations of all the different devices available.

iOS is a closed source operating system developed by Apple. It is designed exclusively for the iPhone and support iPod Touch, iPad, iPad mini and Apple TV. It is developed in C, C++, Objective-C, and Swift, and includes technologies as OS X kernel, BSD sockets for networking and Objective-C and C/C++ compilers for native performance. Also this operating system has a wide-range of graphics capabilities available through high level frameworks.

Apple products generate high profits and the applications developed for their products are the most profitable. A key advantage for the iOS developers is the controlled environment of this operating system, where everything is standardized. To sell and distribute applications to Apple the developers need to enroll the iOS Developer Program. And to have an application available on the App Store, it must be approved by Apple first.

### 2.4.1.1 Software

OpenCV (Open Source Computer Vision) is a large and complex library of C/C++ programming, originally developed by Intel, for image processing and analysis operations. It can be used under a BSD license for academic and commercial purposes. It is cross-platform and runs on Windows, Linux, Mac OS, iOS and Android. The main focus of this library is real-time image processing and computer vision ranging from interactive art to advanced robotics. Some of the features included are image data manipulation, image input/output, basic image processing, structural analysis, object recognition, basic GUI and image labelling. Although it is written in C++ and C, there are interfaces to other languages like Python, Java and MATLAB/OCTAVE [[Aga06](#), [GR03](#)].

MATLAB (MAtrixLABoratory) is a high-level language and interactive environment, developed by MathWorks and released in 1984. It is closed source and can only be used under a property commercial license, with a cost of \$2150. MATLAB is used by millions in academic and research institutions and industrial enterprises. It is a technology that crosses many areas like signal and image processing, communications, control systems and computational finance. Although MATLAB provides an Image Processing Toolbox, most operations can be performed without this toolbox, since it provides a wide number of operations, functions and algorithms in the area of image processing.

## 2.4.2 Preprocessing algorithms

Image preprocessing can increase significantly the reliability of digital image operations, enhancing the visual appearance of an image and improving further processing of the images. Therefore preprocessing techniques are widely used. A brief summary of the most common image preprocessing operations is presented in the following subsections.

### 2.4.2.1 Red-Green-Blue (RGB) to grayscale conversion

RGB is a popular colour model and has been used in various applications [Rui97]. An image represented in the RGB colour space is composed of three independent channels for each colour red, green and blue. The conversion of that image to grayscale only implies the usage of one channel colour space. Although at the user's point of view, images represented in RGB colour space provide more detailed information, certain operations do not require such complex representation, simplifying the algorithm and reducing the computational effort. The conversion to grayscale is done using a weighted average of the RGB values, resulting in the luminance of the inverted color.

### 2.4.2.2 Binarization

Binarization converts a multi-levelled image into a binary image, with two levels of intensity, black and white [GJG07, SD01]. Each pixel is compared with a threshold value. An adequate threshold helps dividing an image into different classes, usually background and foreground. The binarization clears the background and enhances some details of the image. It is important to choose a value that preserves the important parts. There are many methods to do this operation and can be divided in fixed and adaptive. A global fixed value for every pixel may not be good in all situations, so in adaptive methods each pixel have a calculated threshold value. Some popular methods are:

- **Global fixed threshold** — The global thresholding algorithm uses a fixed intensity threshold value  $T$  from 0 from 255. If a pixel on an image have an intensity value bigger than this value, the pixel is set to white, otherwise is set to black. In difficult images this method doesn't show great results.
- **Otsu threshold** — Otsu's thresholding method tries to find the threshold value where the sum of variances of two classes (background and foreground) is minimum. It has problems with local illumination. There are also many variations of this method to meet different conditions.
- **Gaussian threshold** — Gaussian method is an adaptive thresholding, in which the "threshold value is the weighted sum of neighbourhood values where weights are a Gaussian window." [Ope15].
- **Niblack's method** — Local thresholding algorithm where the threshold value adapts to the local mean and local standard deviations over a specific window size around each pixel

location. It can fail in large homogeneous areas [SD01] and in images with large variations in illumination [BRHT04].

#### **2.4.2.3 Digital filtering**

Image filtering is the conversion of an image into a new one, pixel-per-pixel. Calculations are conducted not only considering the given pixel, but also the pixel on its neighbourhood. It is a complex and contextual transformation that aims to solve the challenging problems of noise reduction, correcting blurry images with low contrast, removing image faults, enhancing image elements, reconstructing an image and detect edges and corners, between others. There are two groups of filters: linear and non-linear. Linear filters are the combination of the pixel values in the source image, while non-linear filters include logical, statistical and adaptive operations. Although linear filters are simpler, non-linear ones can present better results.

#### **2.4.2.4 Canny Edge Detector**

Edges represent significant local intensity changes that provide some clues of object boundaries in an image. The Canny Edge detection operator was designed by John F. Canny in 1986 to be an optimal edge detector. This operator uses a multi-stage algorithm in which it classifies a pixel as an edge if the gradient magnitude of the pixel is larger than those of pixels at both sides in the direction of maximum intensity change [DG01].

#### **2.4.2.5 Hough Line Transform**

The Hough Line Transform is a technique of image processing that can be used to detect straight lines. This technique detects any shape, if it is possible to represent that shape in a mathematical form. Although this technique is computationally heavy, it can be used even if the image contains noise or some parts removed.

#### **2.4.2.6 Perspective Correction**

“Perspective distortion occurs due to the perspective projection of a 3D scene on a 2D surface” [GKM13]. Images with perspective distortion often decrease the accuracy of image analysis algorithms [JJ05]. To rectify the perspective distortion some important clues can be obtained by the boundaries of documents, page layout and textual structures.

#### **2.4.2.7 Morphological Operations**

Morphological operations are based in the analysis and processing of geometrical structures. These operations apply a small shape or template, named the structuring element, to an input image, creating an output image of the same size. This structuring element is positioned in all the possible locations in the input image and is compared with the neighbourhood pixels in each position. The result is obtained from testing, if it fits, hits or intersects the neighbourhood. The

fundamental morphological operations are: erosion and dilation. These operations are used for noise removal, and isolation or joining elements in an image.

### 2.4.3 Optical Recognition Character and Software

Optical Character Recognition (OCR) is a field of computer science that classifies optical patterns that correspond to characters. OCR works by converting printed text images or handwritten text into ASCII text which can be later stored and manipulated by a computer [PPP12]. Although the first occurrence of OCR technology was in 1929, the first commercial product was only introduced in 1978, by Kurzweil Computer Products. This technology allows machines to recognize text automatically by analysing source documents and performing character analysis of the resulting images, and then be stored and electronically manipulated (edited, written and displayed) like any standard document. With electronic media gaining importance the need to transfer old documents to electronic representation is growing. It is a leading technology in the software market and very useful [BGR07, LS99]. There are many applications of using OCR, like coding libraries, detecting license plates, translating text, keywords searching, between others, and can be used in many areas as management, business, criminal and social networks [AAAA14] The popularity of this field [DTJT96, TMBY13] lead to a lot of development of many different of techniques, tools and engines [DTJT96]. There are many OCR engines available on the market. Although the majority is expensive and closed source, free open source engines can be found. Generally, commercial OCRs are a best option, support more languages, have complex user interfaces and have a better accuracy in situations where the image quality is low [Smi07b]. But adding some image processing operations to improve the quality of the source images and training the algorithm can put the open source products to the same level of accuracy.

A typical OCR system performs in several steps:

- **Optical scanning** — Where the digital image is acquired, usually by a scanner or a camera.
- **Location and segmentation** —Usually OCR engines isolate characters and perform the recognition individually. Common problems in this step are character fragmentation, distinguishing noise from the text, mistakenly taking graphics or geometry for text, and confusing text for graphics or geometry.
- **Preprocessing** — use image processing techniques to improve the quality the image and increase the chances for success.
- **Feature extraction** — extract the most relevant information of the image that can uniquely identify the text segment. There are a lot of techniques and each system should use a method that best suit their needs [DTJT96].
- **Classification** —matches the image with pre-defined classes.
- **Post processing** —improve the recognition of the text segment. Some techniques used include grouping, error detection and correction techniques.

The machines can't compete with human reading capabilities yet [TMBY13]. Although the accuracy of this technique is generally high [PPP12] and in most of tasks is up to 95%. The performance is directly proportional to the quality of the input image [PPP12, PCAAL00, TMBY13].

Some common situations leading to a decrease of the quality OCR, often to no quality, include:

- **Dynamic backgrounds** — Text is sometimes printed against shaded or textured backgrounds or is embedded in images, adding a lot of noise to the images. The accuracy of extracting digital text in clean images is very high, but when it comes with shaded and textured background this accuracy decreases a lot [NESM01].
- **Text skew** — Images obtained through cameras usually have some distortion due to perspective and warping. Current techniques do not handle well these situations and additional processing needs to be done to restore the frontal view [LDD05].
- **Lighting variations and blur** — Presence of lighting variation (shadow, reflection, flooding, etc.) and blur are usual in images, especially taken by a camera, and can add a lot of noise to the images and even make the text unreadable. Some image processing techniques [FSS13], can be used to recover some of the text but it is not guaranteed to be successful [SD01].
- **Low resolution** — Cameras and scanner with low resolution present a challenge to OCR because they make a good segmentation almost impossible [JSVR05, NESM01]. A possible solution is training the OCR system.
- **Non-standard fonts** — OCR engines use a list of pre-defined fonts to have a better performance, presenting some problems when text is presented in non-standard fonts, although nowadays most engines have a high accuracy even if it is not a common font.
- **Degradation/Aging** — A frequent problem in degraded sources are varying kerning (space between letters), varying leading (space between lines), frequent line-break hyphenation, etc.

Some solutions or improvements to the presented problems include: training the OCR system where an algorithm can be trained on a data set of known characters and can improve the process of classification; cropping the area of the image where the text is; adding a dictionary that guarantees that the words generated are correct and belong to a set of valid words [PCAAL00]. For each problem there is more than a possible solution that needs to be adapted to the specific needs and constraints.

#### 2.4.3.1 Tesseract

Tesseract is an open source OCR engine that can read a wide variety of images in over 60 languages. It was developed at HP between 1985 and 1994 [PPP12, Smi07a], as a PhD research to

create an OCR function in desktop scanners. Since 2005, Tesseract has been maintained and extensively improved by Google. Originally, Tesseract was designed to recognize only English text. Currently is being modified to be able to deal with other languages and other characters besides the Latin ones. It is highly portable [PPP12] because it is an engine and not a tool, which means that it is possible to execute it inside any application. It can run on Linux, Mac, and Windows, etc. The fact of being free and open source, highly portable and having high accuracy are the main reasons of being one of the most popular tools for OCR [KLM09]. Tesseract follows a step by step architecture as shown on Figure 2.10 [PPP12].

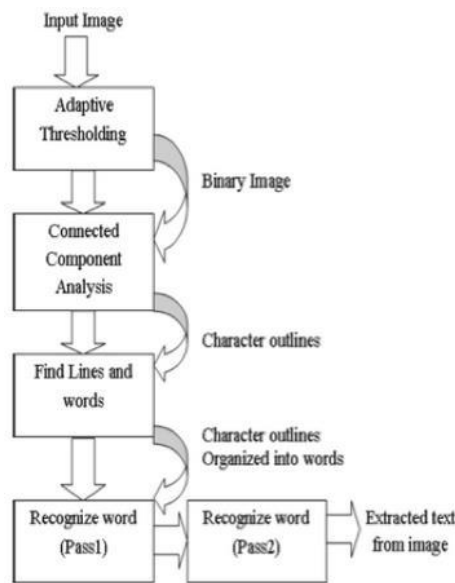


Figure 2.10: Tesseract Architecture. Image taken from [PPP12].

After the three first steps, the text is broken into words. It then proceeds to the recognition of each word as a two-pass process. In the first pass, an attempt of recognition is made. If a word is satisfactory it is passed to an adaptive classifier as training data. The adaptive classifier then gets a chance to more accurately recognize the words that didn't pass [Smi07a]. When problems arise, many solutions are recommended to overcome them. Adding some preprocessing techniques to enhance the source images is the most usual, solving the majority of the problems.

## Chapter 3

# Technical Requirements and Architecture

This chapter focus on the presentation of the technical requirements of the project. Some details of the application, including objectives, features, use cases and requirements are initially presented. Then, the system architecture and the explanation for the choices regarding the technologies used, are shown. The packages diagram of the mobile application is also described in this chapter.

### 3.1 Formal presentation of the objectives, features, use cases and requirements of the project

The nutrition labels provide useful information about a food product to the consumer. By checking these labels regularly, the consumer can be more conscious of the contribution of each product in the overall diet, and consequently make healthier choices. The main objective of this project is to develop a tool that helps the user to analyse relevant nutritional information from a product, by means of a mobile application. This application allows the user to take a photo of the nutrition declaration present in a product and automatically extracts the nutritional information. Only the nutritional declarations are analysed, considering that this type of labels presents the most detailed information. This type of labels can be found in the back of a product and are usually displayed in a tabular form, as a panel or a grid, with the nutrient names and the respective values aligned.

After the analysis of the consumer response to the use of the nutrition labels, described in the Section [2.2.2](#), it is clear the necessity of a solution to facilitate the reading and analysis of the nutrition labels. To meet this need, the application should be easy to use and should display the relevant information on a single, cross-sectional shape. It should also follow the new rules defined in the Regulation (EU) 1169/2011 that sets several rules including that the amount of each nutrient/energy need to be expressed using absolute values and listed in a certain order. It was concluded from the analysis of the studies, presented in the Section [2.2.2](#), that the consumers like

the idea of the FoP labels. Considering this, the complex and detailed information presented, in the nutrition declaration, is complemented with a summary of the nutritional information, like is usual in all the FoP labels.

Accordingly to the bibliography presented in the Section 2.2.2, the consumers find inconvenient and time consuming to compare labels with different formats. In addition it was also concluded that the consumers are able to understand certain key terms and apply this information in simple tasks like comparing nutritional characteristics of products. However, assessing which product is a healthier choice have more factors to consider: the role of the product to the diet and the overall nutrients. Considering these problems, the application also allows the user to compare two products of the same category and check which product is the better option.

In order to choose the categories that were included in this project, it was taken into consideration that the food products that the consumers take more attention to the nutritional information are the breakfast products. Therefore, three categories of breakfast products were chosen: Yogurts, Drink Yogurts and Cookies.

Additionally, the nutrition labels are an important tool to make healthier choices, they can help the consumer to prevent some diseases, and more important its treatments. So, considering any personal data of the user (anthropometric data, diseases, diet restrictions), the user can assess if a specific product can be suitable for his diet.

Taking into account the objectives of this project, the mobile application should then include the following features:

- Take picture of a food label using a mobile device
- Extract and present the nutritional information of a food product
- Compare nutritional values between two products of the same type
- Help users to assess the suitability of a given product, according to their personal data

Use cases are used to provide an user-centric insight of the features, usually divided by actors. However, in this application there is only one actor, the application user. In the Table 3.1 all the use cases are described.

## Technical Requirements and Architecture

Table 3.1: Use cases

<b>ID</b>	<b>Title</b>	<b>Description</b>
US01	Check nutritional information from photo	The user can check the nutritional information of a product by taking a photography of the nutrition declaration of a product
US02	Compare two products	The user can compare two products of the same type by checking the nutritional information of each one
US03	Edit nutritional information	The user can edit the extracted nutritional information of a product
US04	Add personal data	The user can add personal data
US05	Assess the suitability	The user can assess the suitability of a given product, according to his personal data

In order to use this application, it is necessary to take in to account some conditions and needs. These requirements are presented in the Table 3.2.

Table 3.2: Requirements

<b>ID</b>	<b>Title</b>	<b>Description</b>
REQ01	Camera	The smartphone must have a mid to high quality camera
REQ02	Data acquisition	The photo of the label must be focused and in a close-up plan
REQ03	Android API	The API of the Android should be higher than 19
REQ04	OpenCV Manager app	The application OpenCV manager, which is an Android service containing most of the image processing operations, needs to be installed in the smartphone

### 3.2 Architecture and Technologies

The system is composed by two components: a mobile application and a server application, as displayed in Figure 3.1.

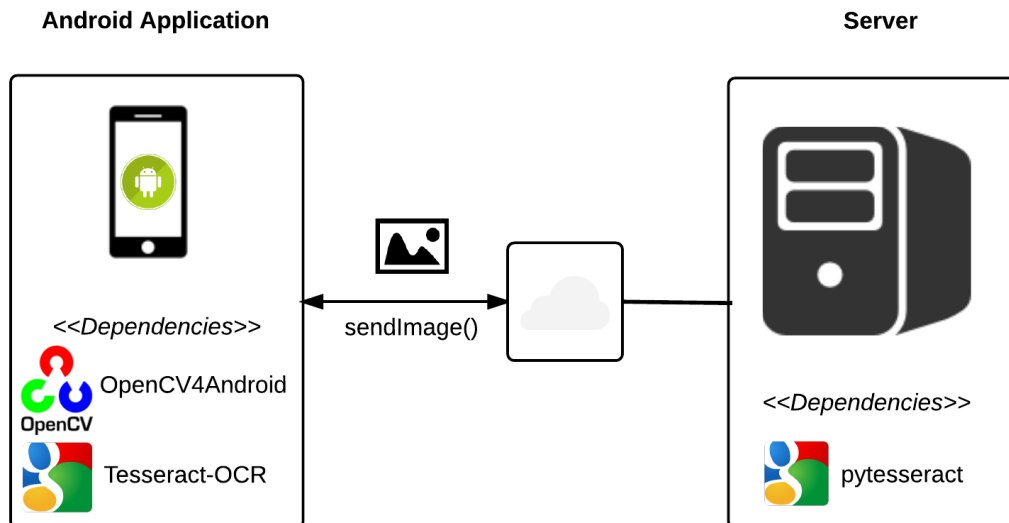


Figure 3.1: System Architecture

The mobile application is the main component of the system and it is developed for Android smartphones. Considering the data collected, the operating system that best suits this problem is Android. Android offers many advantages compared to iOS: freedom to develop any application with thousands of libraries to complement and unlike iOS applications, Android applications can be developed using a popular language, Java, and developed using a free SDK. Also mobile devices that have Android OS are inserted in a larger market, with a big diversity of users with different needs and devices that can meet specific requirements. Some of the needed requirements for this problem include having higher resolution, better cameras and longer durability.

The Android device also uses the OpenCV library for all the operations regarding image processing. As referred previously, both MATLAB and OpenCV are good tools for developing image processing operations. Although MATLAB offers an interactive environment and does not require a solid programming language, it has as main drawback that it is paid. OpenCV, on the other side, can offer many features with usually a better performance with the advantage of being free. Despite MATLAB is not supported yet in mobile devices, it is possible to run it directly in an Android application but it will require additional work.

On the other hand, OpenCV has available a port for Android named OpenCV4Android that offers almost all the features of OpenCV and it is easier to use and integrate. To perform the

conversion from image to text, Tesseract was the OCR engine chosen. Mainly because it is the only free and open source found and, supposedly, presents good results.

The server was developed to complement the Android application. The main objective is to use the server whenever the user has an Internet connection, reducing the time of processing of the conversion of images to text. Since it was not included in the initial planning, it wasn't made an extensive study of the possible technologies. Therefore, was chosen the same OCR engine than the one used in the Android application, Tesseract, and developed in a language supported by this engine, Python. It was installed on a local Linux-server.

The communication between these two components is established through HTTP requests.

### 3.2.1 Packages Diagram

The application is composed by the following packages: Camera, Connection, Feature, Processing, Model, Test. The packages architecture follows the Model-view-controller software pattern, excluding the Test package that is only used to evaluate the system, as displayed in the Figure 3.2.

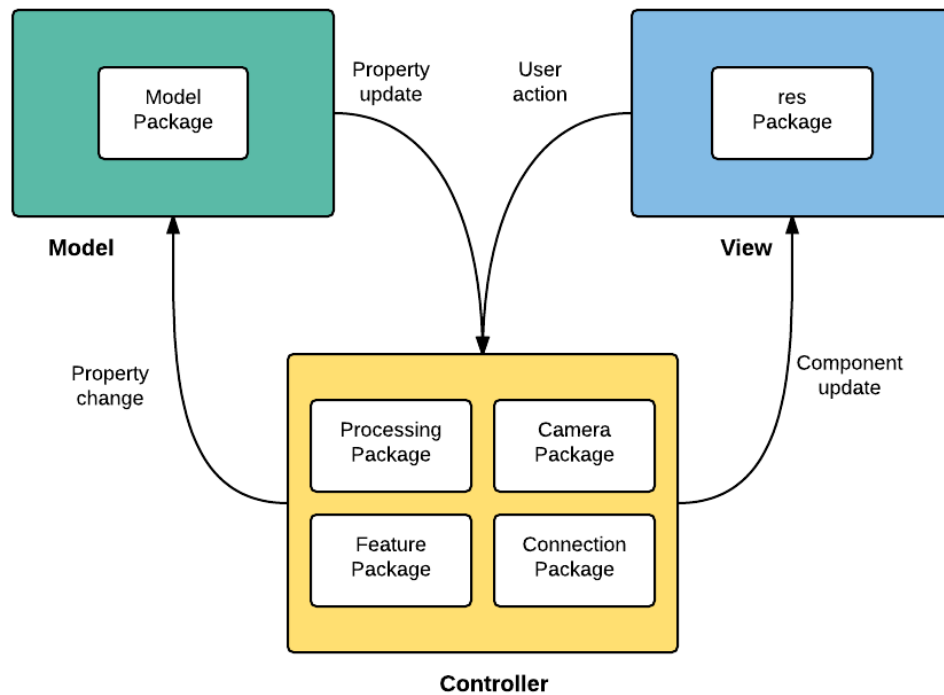


Figure 3.2: Software architectural pattern

• **Controller Package**

The controller packages are, like the name says, the controllers and they are responsible for all the data manipulation. It includes the Camera, Processing, Connection and Feature packages.

In the Camera Package it is defined the behaviour of the camera view and some settings, including the focus of the camera and the acquisition of the frames.

The Processing Package includes all the image processing algorithms including preprocessing operations, OCR and post processing.

The Connection package is responsible for establishing the connection with the server, performing the requests and managing the responses.

The Feature package manages all the changes necessary to the interfaces and user inputs, including all the user interactions regarding the main features.

• **Model Package**

In this module is defined the model class used in this application. The model is displayed in the Figure 3.3.

The *NutritionalInformation* class represents the nutritional data. It is included the nutrients and its corresponding values. Considering that all nutrients have the same attributes and operations, differing only in the values, each nutrient inherits all the characteristics of the abstract class *Nutrients*. The nutrient's classes considered are: *Energy*, *Protein*, *Carbohydrate*, *Sugars*, *Fats*, *Saturates*, *Fibre* and *Salt*. For each nutrient, besides its value, it is also included the reference intake value and the limit values to map to traffic lights color schema.

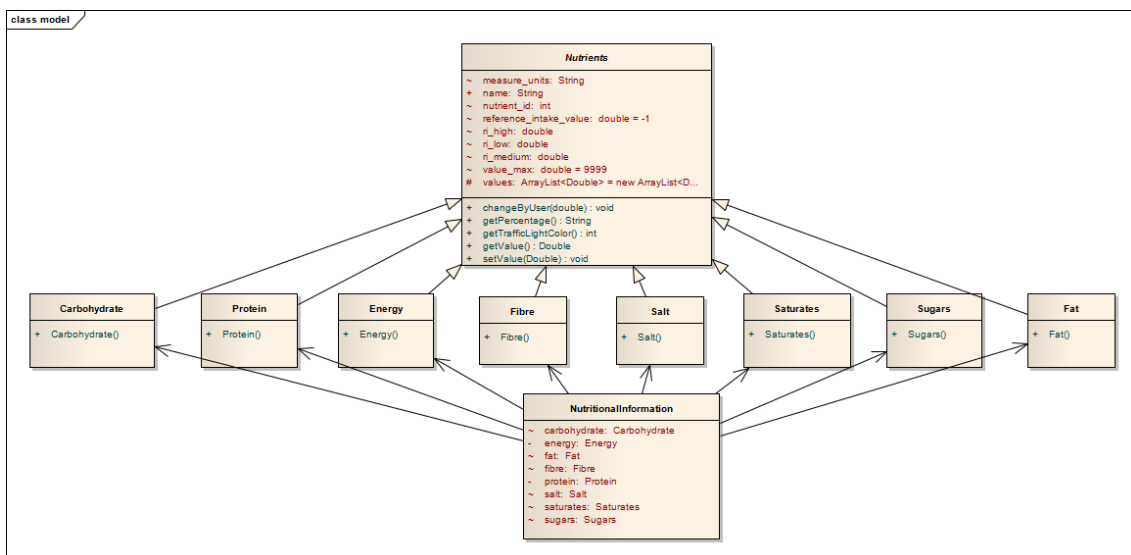


Figure 3.3: Model Class Diagram

## Technical Requirements and Architecture

- **View Package**

The view packages are composed by all the files that are responsible for the interfaces of the application. These interfaces are updated accordingly to the user inputs processed by the Feature package.

These files are included in the default package of resources of all the Android projects, the `android.content.res`.

- **Test Package**

In this package, all the tests that evaluate and validate the processing operations are included.

## Technical Requirements and Architecture

## Chapter 4

# Processing implementation and operation

In this chapter all the image processing operations developed for this project are described in detail. Furthermore, the manipulation of the obtained nutritional information from the image is described. For last, it is shown the interfaces and workflow of the system components.

### 4.1 Image Processing algorithms

The accuracy of the OCR depends largely on the quality of the input image. Due to the fact that the cameras of the mobile device often produce low quality images (blurred, noised, skewed, among others), some types of preprocessing and post-processing techniques were added as an attempt to improve the results.

The user plays an important role in the results obtained. The photo must be taken carefully, trying to reduce to the minimum the amount of reflections, shades and perspective distortions. The photo should also be taken in close up plan, where the label should fill as much as possible the space available on the screen.

It is important to refer that the OCR is a resource consuming process, even without the image processing operations. Causing not only a battery drain, but also a long waiting time for the user, which are factors to be considered in the development of the algorithms.

During the time of the development of this project, most of the nutritional declarations didn't follow the regulation. Consequently, some restrictions were made to the labels that the application can process. The nutrition declaration must have a tabular form, rather displayed as a grid, and all the nutrients must be in the same column as well as the corresponded values. But even by restricting to this type of labels, the information doesn't always present the same order regarding the nutrients neither the same designation of the same nutrient. Besides, they can be displayed

with different font sizes and the type of material where the label is printed also influences the results of the OCR.

This means that even with some restrictions, the operations performed should be the less specific possible in order to be able to work with a larger number of cases.

The algorithm sequence that are present in this document were developed throughout the development of this project, in order to further improve the final results. These algorithms have in consideration the studies presented in the Section [2.4.3](#).

Considering that reducing the area of the image to be processed could improve the results of the OCR and also exclude some erroneous results, the first attempt was to extract the nutritional label of a photo as a preprocessing operation. This attempt deals with the text skew problem. The skew is when the lines of the text are not correctly aligned. This problem causes a decrease of the OCR results, so if the extracted label presents distortion in the perspective view, some geometrical operations are applied in order to restore the frontal view, aligning the text. Two different approaches were developed: the nutrition declaration extraction and the warp of the image, described in the Section [4.1.2](#) and [4.1.3](#) respectively.

Another attempt was to restrict even more the image. The idea is to divide the nutrition declaration for each line of text, as presented in the Section [4.1.4](#).

Taking into account that the OCR performs segmentation and isolation of the characters, another idea was to highlight the location of the words, using a border, in order to facilitate the future location of these characters by the OCR, as detailed in the Section [4.1.5](#).

After discovering that restricting the set of characters of the OCR could be a solution for better results, the operation to divide the image by the column corresponding to the nutrient names and the one that contains the values was also developed, as described in the Section [4.1.6](#).

As concluded in the analysis of the OCR engines, shades and textured backgrounds reduces the accuracy of the OCR. Because of that, the next attempt was to separate the text from the background and remove some of the noise. This attempt is detailed in the Section [4.1.7](#).

Another solution to improve the results is to correct the orientation of the image. Considering the large number of horizontal lines presented on the images, the detection of the angles of these lines can help to calculate the rotation necessary to restore the text alignment, as detailed in the Section [4.1.8](#).

When the text is retrieved from the OCR, a set of regular expression operations is performed to extract the nutrient names and its corresponding values, as described in the Section [4.1.9](#).

### **4.1.1 Extract vertical and horizontal lines**

As the nutritional declarations are usually displayed in a tabular form, with well defined limits, the detection and location of horizontal and vertical lines is widely used in the preprocessing operations. To detect the horizontal lines, morphological operations are performed, using a predefined

structuring element known as the kernel. By choosing the correct size and shape, it is possible to construct a kernel that is sensitive to the specific shape, in this case, the horizontal and vertical lines. In both cases it is applied an adaptive threshold using the Gaussian method to reduce some noise presented in the images. In order to detect the vertical lines presented in the image, the structuring element used in the morphological operations is a vertical line with a relative height. The height of the kernel is always relative to the number of rows of the image and should be smaller or equal than the original image, considering that the vertical lines are within the limits of the image. Depending on the purpose of the operation, different values can be used. The same happens to the detection of the horizontal lines. The difference is that the structuring element used in the morphological operations is a horizontal line. In this case, the width of the kernel is relative to the number of columns of the image.

#### 4.1.2 Nutrition declaration extraction

A possible solution to improve the results is to separate the regions of the image that needs to be analyzed by the OCR. Thereby, the area to compute is reduced, decreasing the computational effort and consequently the user waiting time. It also excludes results that can be obtained erroneously. The nutritional information located in the nutrition declaration is the only data that needs to be extracted. The fact that the declaration is presented in a tabular form, often as a grid, facilitates the identification of the area.

To help identifying the area of interest, a rectangular border is added to the camera view. When the user is taking the photo, the nutrition declaration must be within the limits. This border also ensures that the nutrition declaration is in close up plan and focused. The nutrition declaration needs to fill at least one third of the image to be detected.

To achieve better results, a few preprocessing techniques are sequentially performed on the image. The image is initially converted to grayscale, then an edge detector is used as shown in the Figure 4.1. At last, a closing operation is performed to find and remove potential holes between edge segments.

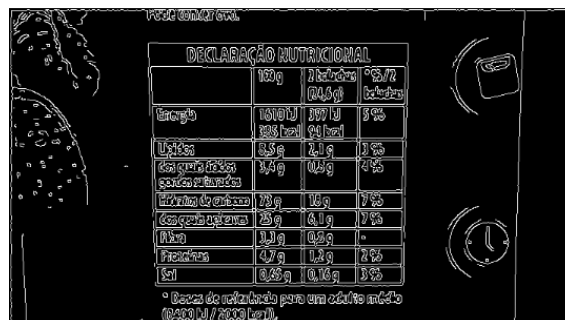


Figure 4.1: Dilation

The table usually presents an outside border, or contour, therefore in order to detect the nutrition declaration, the contours of the image must be found. For each contour the bounding rectangle

is calculated and if the area corresponds to the limit defined, the bounding rectangle is cropped from the original image. The result of this process is presented in the Figure 4.2.

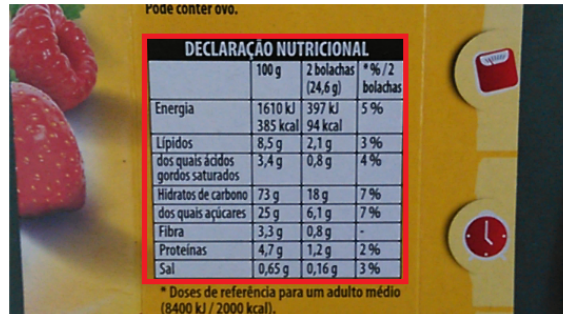


Figure 4.2: Resultant image

### 4.1.3 Warp Image

When the user takes a picture, the result often has perspective distortion. Therefore, the distortion should be corrected before the following analysis. As the nutritional labels have a quadrilateral shape, discovering the 4 corners of the nutrition declaration and obtaining the encompassing rectangle can be used to correct perspective and also reduce the area of search. The first step is to detect the vertical and horizontal lines presented in the image, using the filters described in the Section 4.1.1. The following step is to check the existence of intersection points between each vertical line and horizontal line. An example is presented in the Figure 4.3 where the red dots correspond to the intersecting points found.



Figure 4.3: Intersecting points

After all the points detected it is necessary to find which of them correspond to the corners. The point that has the minimum x and the maximum y is the left top corner, the point that has the maximum y and the maximum x is the top right corner, the point that has the minimum x and the minimum y is the bottom left corner and the point that has the maximum x and the minimum y is the bottom right corner. In Figure 4.4, the corners with the previously stated rules, are shown.

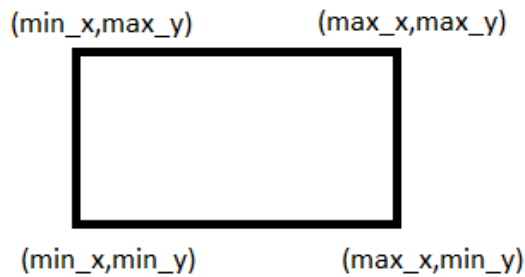


Figure 4.4: Mapping the Points to rectangle corners

After the resultant rectangle is warped, the nutritional label with the correct perspective can be obtained. An example of the resultant image is presented in the Figure 4.5.

Valores nutricionais médios por 100g	Cada porção de 30g de Estrelas + 125ml de leite meio-gordo contém		% VDR*
Valor energético 1645 kJ/388 kcal	739 kJ/175 kcal		9 %
Proteínas 6,1 g	5,9 g		12 %
Hidratos de Carbono 85,0 g	32 g		12 %
dos quais açúcares 29,9 g	15,1 g		17 %
Lípidos 2,0 g	2,6 g		4 %
dos quais saturados 1,3 g	1,5 g		7 %
Fibra 2,8 g	0,8 g		3 %
Sódio 0,3 g	0,1 g		4 %

Figure 4.5: Nutritional label extracted with perspective correction

#### 4.1.4 Divide the Image horizontally

Dividing the image by text lines before passing it to the Tesseract OCR, reduces the area that the OCR needs to analyze. In order to separate the image by horizontal lines, the filter described in the Section 4.1.1 is used. This filter returns a binary image, in which the white pixels identifies the horizontal lines. The Figure 4.6 displays the result of this operation.



Figure 4.6: Separate horizontally processing

It is taken, that the rows with the highest number of white pixels are likely to be the separation between the text lines. For each row of the image, the number of white pixels is counted and stored. Considering that one quarter of the image is composed by horizontal lines, can be assumed that the rows that have a higher number of white pixels than the average of the 25% of the total image are interpreted as line breaks and the image is cropped by those values. Figure 4.7 displays the position of the detected text lines separation.

DECLARAÇÃO NUTRICIONAL			
	Por 100 g	Por 3 bolachas (30 g)	%* Por 3 bolachas (30 g)
Energia	1765 kJ 420 kcal	530 kJ 126 kcal	6 %
Lípidos	12 g	3,6 g	5 %
dos quais ácidos gordos saturados	2,9 g	0,9 g	5 %
Hidratos de Carbono	62 g	18,5 g	7 %
dos quais açúcares	15,5 g	4,7 g	5 %
Fibra	8,6 g	2,6 g	-
Proteínas	11,5 g	3,5 g	7 %
Sal	0,53 g	0,16 g	3 %
Cálcio	320 mg 39 % VRN**	95 mg 12 % VRN**	
Ferro	4,4 mg 32 % VRN**	1,3 mg 9 % VRN**	

Figure 4.7: Separate horizontally after

#### 4.1.5 Text location

Highlighting the specific regions that contain text in the image can help the OCR process. For this purpose, a border was added to the regions that are likely to contain words or group of words. To enhance and help the location of the words, a rectangular Sobel filter and a closing operation with a rectangular kernel were used, since words and groups of words have a rectangular shape. In addition, the horizontal lines are deleted to facilitate the identification of the text lines. In the Figure 4.8, an example of the resultant image, can be seen.

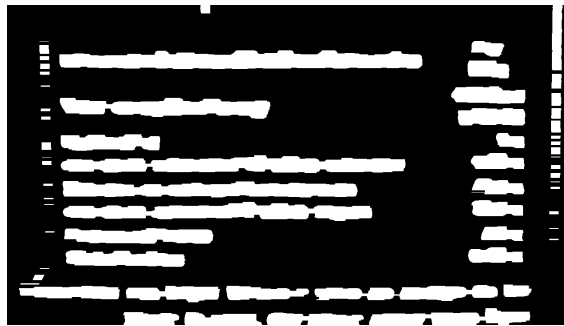


Figure 4.8: After processing

After the contours of the white regions of the image are found, a border is drawn to all contours which are likely to be words or group of words, as shown in the Figure 4.9.

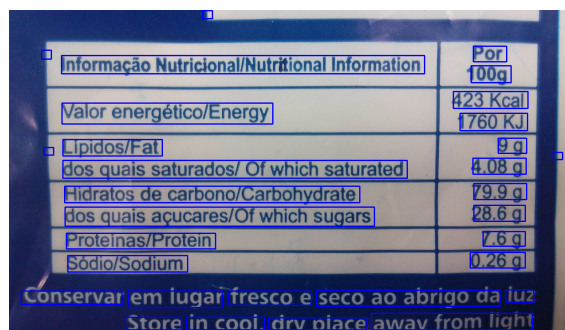


Figure 4.9: Added border to text

#### 4.1.6 Divide the Image vertically

The Tesseract OCR offers the option to create a white list that narrows the set of characters which the OCR will need to recognize, and consequently improve the accuracy. Usually, as the nutrition declaration has a tabular form, and the first column usually refers to the nutrient names, the first column only includes letters. The following columns to the right have the values of each nutrient, including only digits and the respective units (example in Figure 4.11). Dividing the table using the grid separations allows us to obtain the nutrients part and the corresponding values part. The filter described in the Section 4.1.1 can be applied so that the vertical separations remain in the image, as presented in the Figure 4.10. The result is a binary image. The white pixels correspond to the foreground, which is composed by the grid of the nutrition declaration, and the black pixels correspond to the background.

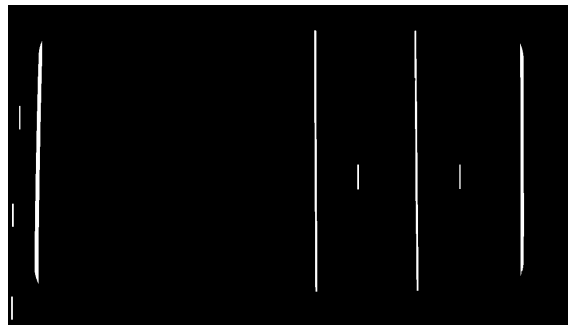


Figure 4.10: Separate vertically

The image columns which have a higher number of white pixels are likely to be the separations. These separations are usually distant from each other so it is not needed to go through all the columns. To locate these separations, for each 10 columns of the image, the number of white pixels is counted and stored in a list. Admitting that about 15% of the image is vertical lines, can be assumed that the columns that have values higher than the average 15% highest values, are likely to be separations. Finally the image is cropped by the separations, as presented in the Figure 4.11.

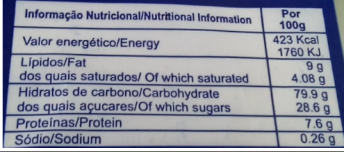
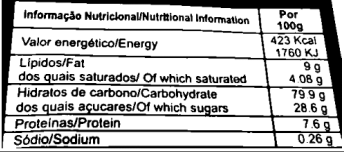
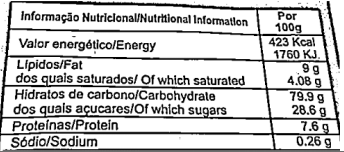
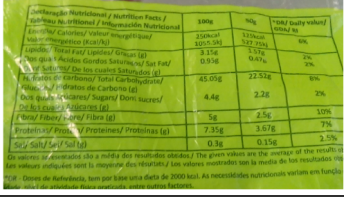
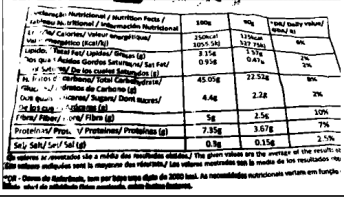
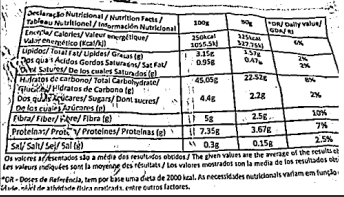
INFORMAÇÃO NUTRICIONAL	Por 100 g	Por 23,5g
Valor energético	1576 kJ 373 kcal	370 kJ 88 kcal
Lípidos	7,0 g	1,6 g
dos quais ácidos gordos saturados	3,3 g	0,8 g
Hidratos de carbono	69,7 g	16,4 g
dos quais açúcares	24,6 g	5,8 g
Fibra	5,0 g	1,2 g
Proteínas	5,4 g	1,3 g
Sal	0,95 g	0,22 g

Figure 4.11: Separate vertically

#### 4.1.7 Enhance the Image

Another possible way to improve the results obtained by the Tesseract is to perform a binarization of the image. This process converts an image to a two-level color scheme. The objective is to clear the background and enhance the text. It is important to try to preserve all parts of the text. The two techniques that obtained best results were the Otsu binarization and the Adaptive Gaussian with Gaussian blur. Despite the Otsu method presents better results when the image is in good conditions and have an uniform illumination, the Adaptive Gaussian is able to return better results when the image is in bad conditions.

Table 4.1: Enhancements

Original	Otsu	Gaussian
		
		

### 4.1.8 Correct orientation

The image might be slightly rotated during the acquisition, which means that the orientation of the text lines is not aligned with the image edges and it causes a decrease of the Tesseract accuracy. Therefore it is necessary to correct the image before going further in the processing pipeline.

The horizontal lines are detected using the Hough Line Transform. Before applying the Transform, some preprocessing operations are recommended. These preprocessing operations are composed by a conversion to grayscale, a smoothing of the image with a Gaussian filter and the application of a threshold by the Otsu method. To ease the detection of the horizontal lines, it is also added a horizontal Sobel filter to enhance the horizontal changes presented in the image.

The original image and resultant image obtained by this processing are displayed in the Figure 4.12 and Figure 4.13.



Valores nutricionais médios por 100g	Cada porção de 30g de Estrelas + 125ml de leite meio-gordo contém	% VDR*
Valor energético 1645 kJ/388 kcal	739 kJ/175 kcal	9%
Proteínas 6,1 g	5,9 g	12%
Hidratos de Carbono dos quais açúcares 85,0 g / 29,9 g	32 g / 15,1 g	12% / 17%
Lípidos dos quais saturados 2,0 g / 1,3 g	2,6 g / 1,5 g	4% / 7%
Fibra 2,8 g	0,8 g	3%
Sódio 0,3 g	0,1 g	4%

Figure 4.12: Original Image



Figure 4.13: Image after processing

After the Transform is applied, the angle of the resultant lines is calculated to ensure that only horizontal lines with a maximum deviation of 10° are considered.

There are two possible rotation directions, clockwise or counter clockwise. In order to correct the image, it is necessary to find each line rotation. In the end, the image is rotated with the average of the rotations, as shown in the Figure 4.14.

Valores nutricionais médios por 100g	Cada porção de 30g de Estrelas + 125ml de leite meio-gordo contém		% VDR*
Valor energético 1645 kJ/388 kcal	739 kJ/175 kcal		9 %
Proteínas 6,1 g	5,9 g		12 %
Hidratos de Carbono 85,0 g	32 g		12 %
dos quais açúcares 29,9 g	15,1 g		17 %
Lípidos 2,0 g	2,6 g		4 %
dos quais saturados 1,3 g	1,5 g		7 %
Fibra 2,8 g	0,8 g		3 %
Sódio 0,3 g	0,1 g		4 %

Figure 4.14: Image after processing

#### 4.1.9 Post-processing

When the results from the OCR are retrieved, the text is filtered by several regular expressions, in order to extract the nutrient names and the values. The first step is to find the nutrient names. A predefined list of the nutrient names is used to construct these regular expressions. As the words often present several errors the regular expressions are constructed to cover many cases becoming more flexible.

#### 4.2 Final algorithm

After the tests performed on the image processing algorithms, it was considered the addition of a server for running the OCR. The time consumed by the server is considerably less and the accuracy is higher.

In order to increase the probability of success, instead of passing one single image taken by the user to the OCR, several previous frames are sent. Considering that the user takes the photo in the best conditions, some of the previous frames are also likely to be close to the better conditions. Taking into account that different processing operations solve different problems, and combined operations didn't bring any improvements, performing one operation in each different frame before sending the frames to the server was the best solution found.

The frame that corresponds to the actual photo, is performed the operation with highest accuracy, which is the enhancements algorithm. In the frame immediately before is used the algorithm to extract the nutrition declaration, the previous frames uses the correction of the orientation. Finally, the last frame which has the highest probability of being in worst conditions, doesn't suffer any operation.

After the OCR is performed in each image, the results are retrieved and all the values are saved in a list. The values with the highest frequency are the ones that will be considered. Additionally, in

order to increase the confidence in the results obtained, some nutrients have a predefined maximum value that is three times higher than the recommended value.

### 4.3 Analysis of the Nutritional Information

When the results are retrieved, the information is presented to the user accordingly to the Regulation (EU) 1169/2011. One of the rules is the order of each the nutrient and its values. As referred in the study of the consumer response presented in the Section 2.2.2, a solution to help the users to have a better understanding of the nutrient values and its impacts is to add some interpretation aids. In this case was used the same idea behind the Traffic Lights, the color coded scheme that presented the better response from the consumer. In addition to the traffic lights, the ratio between the retrieved values and the reference intake values, is shown.

The application allows the possibility for the user to compare two products. Without the specific knowledge, it is very difficult to compare products, considering that each food product can has different functions and roles in the user's diet. This causes that each product has its limits of values but they can be grouped in categories that share the same limits. The three categories chosen were: Yogurts, Drink Yogurts and Cookies. In order to learn which product is better to the user, a nutritionist was consulted. There isn't a formula to compare products, many factors need to be considered in order to assess the best product. The most generic rule that is still reliable is to compare the information of nutritional with products considered "healthier". Therefore, the best option is the one that presents the smallest difference to that product.

The values considered in each category are presented in the tables below.

In Table 4.2, are displayed the healthier values for 100 g of yogurt.

Table 4.2: Healthier values for a yogurt

<b>Nutrients</b>	<b>Values</b>
Energy (kJ)	286.39
Proteins (g)	5.0
Fats (g)	0.1
of which saturates (g)	0.05
Carbohydrates (g)	11.8
of which sugars (g)	10.03
Fibre (g)	0

In the Table 4.3, the healthier values for 100 ml of drink yogurt, are shown.

Table 4.3: Healthier values for a drink yogurt

<b>Nutrients</b>	<b>Values</b>
Energy (kJ)	256.27
Proteins (g)	3.2
Fats (g)	0.3
of which saturates (g)	0.15
Carbohydrates (g)	11.6
of which sugars (g)	9.86
Fibre (g)	0

In Table 4.4, are displayed the healthier values for 100 g of cookies.

Table 4.4: Healthier values for cookies

<b>Nutrients</b>	<b>Values</b>
Energy (kJ)	1825
Proteins (g)	8.4
Fats (g)	12.2
of which saturates (g)	4.88
Carbohydrates (g)	72
of which sugars (g)	14.4
Fibre (g)	0

## 4.4 Mobile Application interfaces and features

The mobile application is aimed to be incorporated in the Smart Companion project, which was specially designed to meet older adults' needs. Therefore the design of the application follows the guidelines defined by Fraunhofer Portugal Aicos. The application allows the user to check the nutritional information on a product and compare two products of the same category.

- **Check nutritional information**

This feature allows the user to check the nutritional information in a more user friendly way. The workflow of this feature is presented in the Figures 4.15 and 4.16. After the user selects "Check Nutritional Information" (a), an option to take a photo is available (b). The user should take a picture of the nutrition declaration of a food product package (c). The majority of the labels has a horizontal shape, the camera view is set to landscape orientation.

## Processing implementation and operation

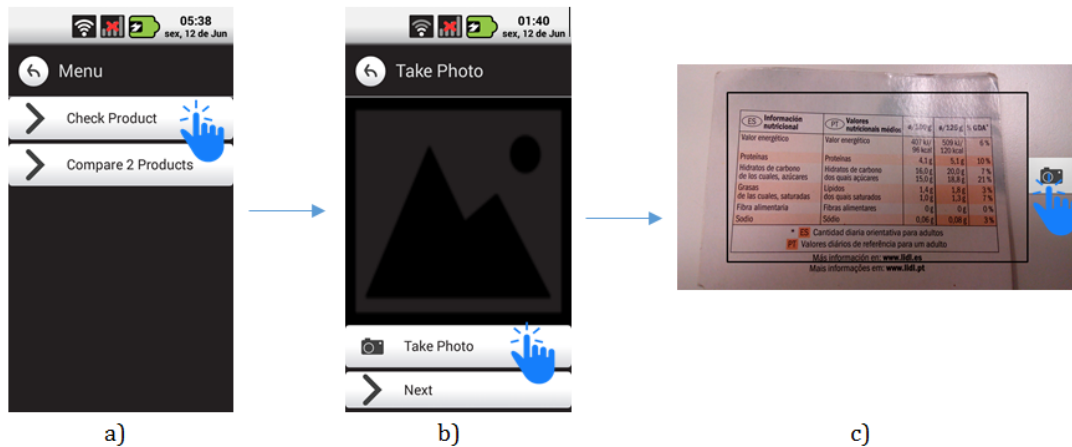


Figure 4.15: Part 1: Workflow to check the nutritional information of a product

The quality of the photo has a large contribution to obtain better results, the user can preview the image and take a new photo if not satisfied. To proceed, the user should select the option "Next" (d). The conversion to text is performed and if any results are retrieved, the nutritional information is displayed (e).

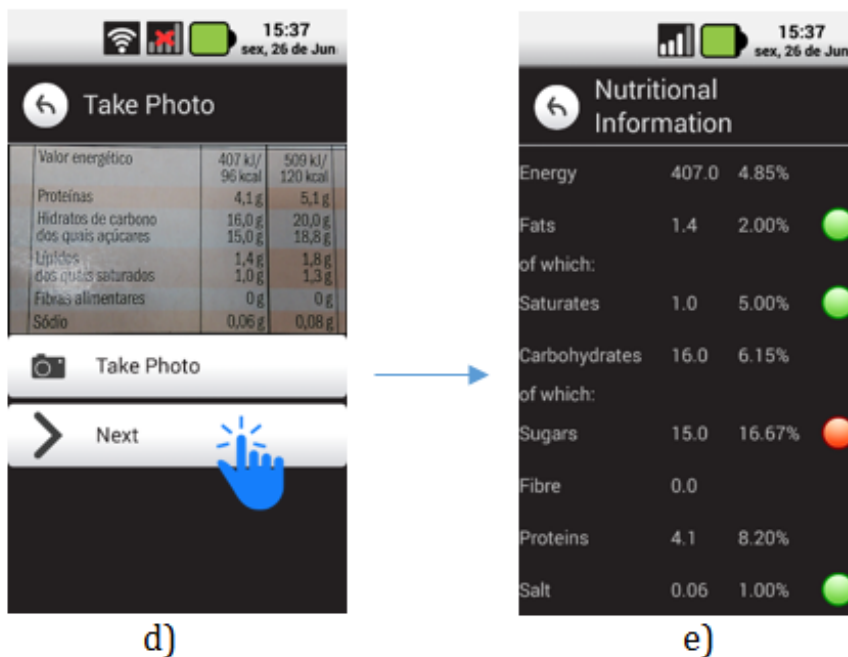


Figure 4.16: Part 2: Workflow to check the nutritional information of a product

Since the results can contain some errors, the user can change the values retrieved following the steps displayed in the Figure 4.17. The user simply needs to click on the values that he wants to change (f). A text box will appear where the user can write the new value

## Processing implementation and operation

and overwrite the previous value by selecting the "Save changes" option (g). Finally, the information will be updated and presented with the performed changes.

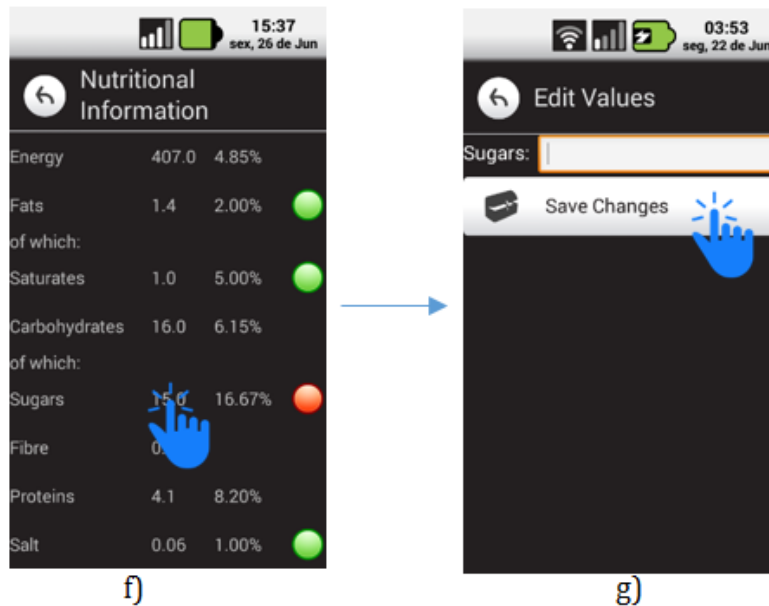


Figure 4.17: Edit the nutritional values

### • Compare two products

This feature helps the user to assess which is the healthier option between two food products of the same category. The workflow of this feature is presented in the Figures 4.18 and 4.19.

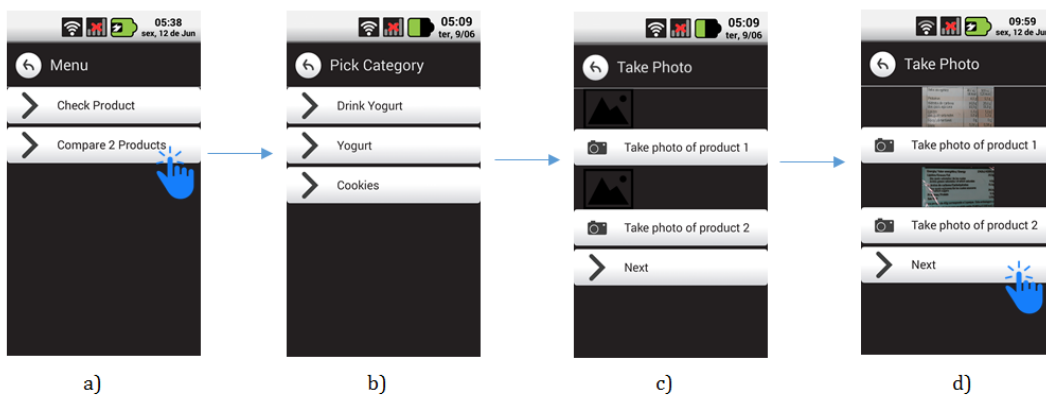


Figure 4.18: Part 1: Workflow to compare two products

In the menu page, the user needs to select "Compare 2 products" (a). A menu with the possible categories appears (b). After selecting food category, the user can take a photo of each nutrition declaration (c). As with the previous feature, the user can preview the images

## Processing implementation and operation

(d) and re-take the photos. To finalize this feature, the user should select the option "Next" (d).

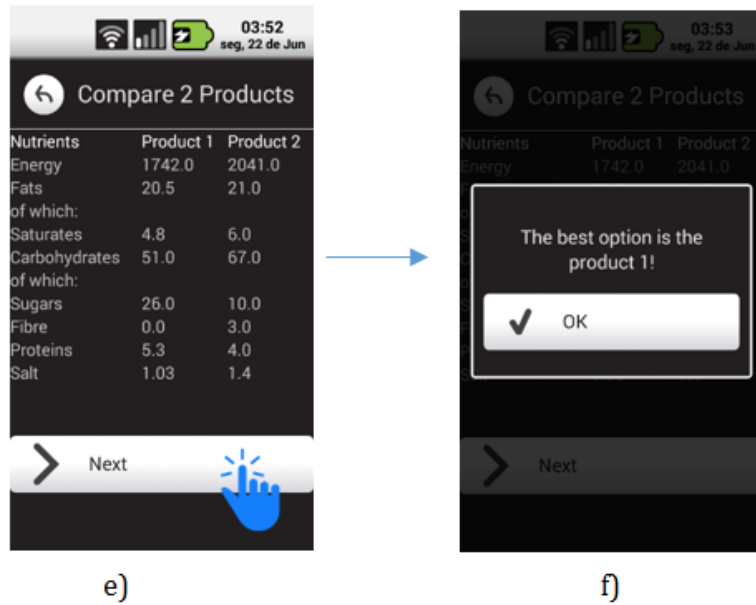


Figure 4.19: Part 2: Workflow to compare two products

The values retrieved from each product are shown using a table (e), so the user can check the differences between the values. A message will be displayed stating which product is the better option (f). As in the previous feature, all the values also can be changed following the same steps.

## Processing implementation and operation

# Chapter 5

## Results

Several tests were made to evaluate the accuracy of some of the developed algorithms and its contribution for obtaining better results. The separation regarding the division of the image, vertically and horizontally, was discarded due to the bad results retrieved. For this purpose five sets of images were taken by five different smartphones. This set of images includes 14 pictures of food labels of the nutritional declarations, where 9 of the images are from the labels of cookies products and 5 are from yogurts. One of the sets is presented in the Annex A. The same set was used for all the tests, including the Android application and the server tests.

The specifications of the smartphones used for the tests are described in the Table 5.1.

Table 5.1: Smartphones specifications

<b>ID</b>	<b>Brand</b>	<b>Model</b>	<b>Camera</b>	<b>Image resolution</b>
Elephone	Elephone	P3000s	13MP	(1280x720) px
GalaxyNexus	Samsung	Galaxy Nexus	5MP	(2592x1944) px
HuaweiG6	Huawei	G6-L11	8MP	(3264x2448) px
MotoG	Motorolla	XT1032	5MP	(2592x1456) px
Vodafone985N	Vodafone	985N	5MP	(2560x1440) px

To perform the tests a Toshiba Satellite P50-B-10H running Windows 8.1 64-bit was used, with a processor Intel Core i7-4700Q, a CPU with a frequency of 2.4GHz and 16GB RAM.

The first test helps understand the need to use an external server to perform the OCR. For this purpose, the accuracy of the OCR in both server and mobile application, without any processing is shown. The average of the consumed time, for each smartphone, is also presented.

## Results

Table 5.2: Android results by smartphone

<b>ID</b>	<b>Accuracy</b>	<b>% 0 matches</b>	<b>Time</b>
Elephone	17.34%	71%	1519ms
GalaxyNexus	26.19%	36%	5925ms
HuaweiG6	14.63%	57%	11510ms
MotoG	20.75%	22%	2194ms
Vodafone985N	21.77%	50%	2007ms
Average	20.14%	53%	4631ms

Table 5.3: Server results by smartphone

<b>ID</b>	<b>Accuracy</b>	<b>% 0 matches</b>	<b>Time</b>
Elephone	30.87%	43%	1117ms
GalaxyNexus	27.68%	57%	3202ms
HuaweiG6	11.61%	71%	2248ms
MotoG	38.39%	36%	2486ms
Vodafone985N	40.05%	21%	1917ms
Average	29.7%	30%	2194ms

As presented in Table 5.2 and 5.3, the server presents more 9% of accuracy and less 23% of 0 matches. In addition it consumes less 47% time.

In order to improve the results obtained by the OCR, many preprocessing techniques are used, as referred in the Section 4.1.

First the operations are performed on the images in the Android application. After, the modified images are passed to the server who will convert them to text and, by using the post processing operation developed, the nutrients and correspondent values are mapped. To calculate the accuracy of each algorithm, a file with the actual values is used to confirm the results obtained. The number of matches between the values found and the actual value is calculated for each set of image of the smartphone and in the end the average of matches for each smartphone is calculated. In this section only the averages by the smartphone are presented. To better understand the values obtained, is also counted every time the image returns 0 matches between the values retrieved and the expected. So the counting of the average of zero matches by smartphone is also presented. To assess the performance, it will be also calculated the time consumed in the process of each image (excluding the time of the OCR that is the same as presented above), but only the average of the times per smartphone will be presented in this section. It is important to refer that two executions of the same image, can return different results.

## Results

The extraction of the nutrition declaration presents an accuracy of 25% with 29% of images with 0 matches. It takes in average 290ms to perform. These results are presented in the Table 5.4. Although the accuracy decrease, the percentage of 0 matches show a little improvement.

Table 5.4: Nutritional declaration extraction tests

ID	Accuracy	% 0 matches	Time
Elephone	22.45%	50.0%	53 ms
GalaxyNexus	28.57%	28.6%	201 ms
HuaweiG6	16.32%	35.7%	443 ms
MotoG	29.59%	28.2%	519 ms
Vodafone985N	28.57%	14.3%	233 ms
Average	25.1%	28.6%	290 ms

Warping the image presents an accuracy of 28.54%, takes in average 852ms to perform and 41% of the images retrieve 0 matches. These values are presented in the Table 5.5. Although the accuracy increase a little, considering the previous operation, the percentage of 0 matches worsens significantly.

Table 5.5: Warp Image tests

ID	Accuracy	% 0 matches	Time
Elephone	33.16%	50%	206 ms
GalaxyNexus	30.36%	50%	896 ms
HuaweiG6	13.65%	57%	1664 ms
MotoG	36.99%	21%	778 ms
Vodafone985N	28.57%	29%	719 ms
Average	28.54%	41%	852 ms

The text location presents an accuracy of 7.14% with 77% of images with 0 results. It takes an average of 336ms to perform, as presented in the Table 5.6. The results are highly unsatisfactory.

Table 5.6: Text location tests

ID	Accuracy	% 0 matches	Time
Elephone	7.14%	78.6%	71 ms
GalaxyNexus	2.04%	85.7%	391 ms
HuaweiG6	14.29%	71.4%	612 ms
MotoG	9.18%	71.4%	299 ms
Vodafone985N	7.96%	78.6%	310 ms
Average	8.12%	77.1%	336 ms

## Results

The image enhancement presents an accuracy of 39.8% with 30% of images with 0 matches. It takes in average 71ms to perform, as presented in the Table 5.7. This was the only processing that shows better results with accuracy increasing, but the percentage of 0 matches didn't show any changes.

Table 5.7: Enhance the Image tests

<b>ID</b>	<b>Accuracy</b>	<b>% 0 matches</b>	<b>Time</b>
Elephone	54.08%	28.6%	17 ms
GalaxyNexus	34.69%	35.7%	82 ms
HuaweiG6	20.40%	42.9%	142 ms
MotoG	47.96%	14.3%	61 ms
Vodafone985N	41.83%	28.6%	63 ms
Average	39.79%	30.0%	71 ms

The correction of the orientation presents an accuracy of 29.4% with 47% of images with 0 matches. It takes in average 301.24ms to perform. These results are presented in the Table 5.8. These operations also show worst results that using the image without any processing.

Table 5.8: Correct orientation tests

<b>ID</b>	<b>Accuracy</b>	<b>% 0 matches</b>	<b>Time</b>
Elephone	34.69%	42.9%	69 ms
GalaxyNexus	28.57%	57.1%	348 ms
HuaweiG6	13.27%	71.4%	561 ms
MotoG	30.61%	42.9%	259 ms
Vodafone985N	39.80%	21.4%	270 ms
Average	29.39%	47.14%	301 ms

## Results

The algorithms retrieved unsatisfactory results, while the best algorithm, which is enhancing the image, retrieves almost 40% of accuracy with 30% of 0 matches. An attempt to improve these results was to combine several algorithms. However, this attempt was excluded since the resultant accuracy was almost the same as the best algorithm used in the combination. The one that retrieved better results was correcting the orientation and performing the enhancements. The results of this combination are shown in the Table 5.9.

Table 5.9: Combining algorithms

<b>ID</b>	<b>Accuracy</b>	<b>% 0 matches</b>
Elephone	47.83%	29%
GalaxyNexus	28.57%	36%
HuaweiG6	21.17%	43%
MotoG	52.17%	14%
Vodafone985N	43.62%	29%
Average	39.9%	30%

Other solution was to send the last four consecutive frames to the server. Without any processing, the results obtained show an improvement in about 10% in comparison to the best approach, as presented in Table 5.10.

Table 5.10: Summary of the averages calculated of the image processing operations

<b>Results</b>	<b>Without preprocessing</b>	<b>Label extraction</b>	<b>Warp Image</b>	<b>Enhancements</b>	<b>Correct Orientation</b>
Accuracy	29.7%	25.1%	28.54%	39.8%	29.4%
0 matches	30%	28.6%	41%	30%	47.14%

## Results

As concluded above, combining the operations doesn't improve the results. Considering that the OCR works better with different algorithms, not combined, but alone, that means that fixing one thing at a time is the best approach. So another attempt was to use the four different images with four different algorithms. Many combinations were tried, and the best one shows an improvement of 15% of the accuracy of the OCR and reducing the 0 matches in 22% compared to the best algorithm, as presented in the Table 5.11. Considering that the time for running this algorithm and the OCR time of the 4 images, the total time consumed in this operation can reach:  $896\text{ms} + 4 \cdot (2194) = 9672\text{ms}$  (depending on the Internet connection speed and mobile device specifications).

Table 5.11: Multiple images tests with different processing

ID	Accuracy	% 0 matches	Time
Elephone	54.21%	7%	253ms
GalaxyNexus	54.85%	7%	1468ms
HuaweiG6	43.88%	21%	261ms
MotoG	63.65%	0%	1088ms
Vodafone985N	58.16%	7 %	1410ms
Average	54.95%	8%	896ms

After evaluating the results of all the tests, the averages of each algorithm can be found in the Figure 5.1. After analysing the graph it is possible to conclude that the best results were obtained using different frames of the same photo and perform a different operation in each one.

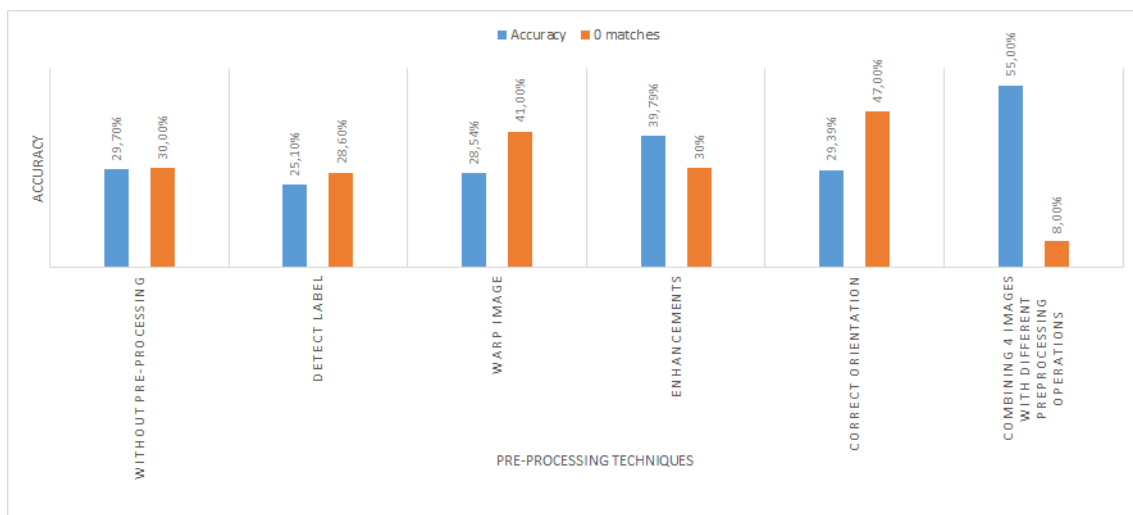


Figure 5.1: Compare accuracy by image resolutions

## Results

In Figure 5.2, is shown the accuracy of some preprocessing operations for the tested smartphones. This analysis allows a linkage between the average of the accuracy and the image resolution, for each smartphone.

The smartphone that shows better results in the overall of the image processing operations is the Moto G, followed by the Vodafone 985N that has the same camera capacity and similar image resolution. Despite expectations, the Huawei G6 presents the worst results. This is an unexpected result because the smartphone has the biggest image resolution and the second best camera capacity. The results shown that the image resolutions are almost inversely proportional to the accuracy. However, the samples used for testing are limited and its conclusions not fully trustworthy.

Another point analyzed is that smartphones with the same image resolutions show similar results, but as previously referred the sample is not enough to make a reliable conclusion.

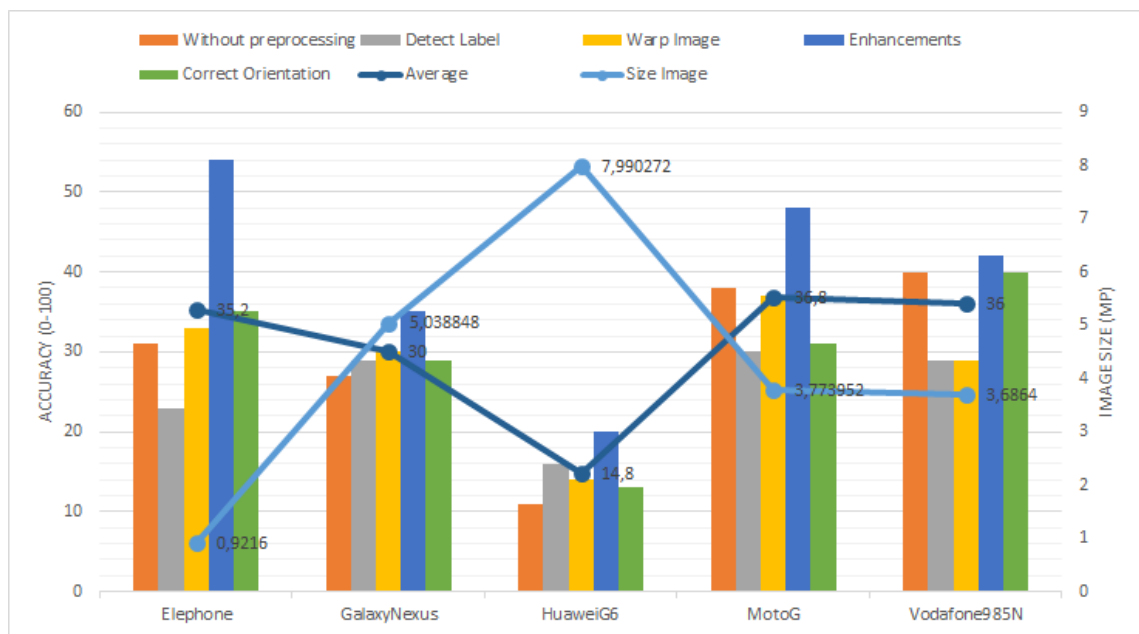


Figure 5.2: Compare accuracy by image resolutions

## Results

## Chapter 6

# Discussion, conclusions and future work

In this chapter, an overview of the implemented system is made, as well as an analysis of the main results, describing the conclusions obtained from it. Future work possibilities are also presented.

### 6.1 Discussion of the results

The nutritional declarations analyzed throughout the development of this project didn't follow the regulation established or have a standard format. Due to the large variety of different nutrition labels, some restrictions were made regarding the labels that the application can process. Nevertheless the restrictions aren't very specific, which leads to a large number of possible cases that the image processing operations must be able to handle. It is necessary to take into account also that smartphones often produce low quality images with perspective distortion, blurry and with noise.

In the initial planning, the application was designed to be standalone. However, due to the unsatisfactory results obtained with different preprocessing operations and a high sensitive post-processing, it was necessary to find a better solution. The possibilities were to continue the image processing technique attempts or to develop a server to perform the OCR operation. As the time consumed by the Android application was already too long and the results of the OCR on the server show an improvement, even without considering the image processing operations, this was the chosen solution.

Although the server obligates the user to have an Internet connection, the Tesseract available for the server resulted in an improvement of more 9% on the accuracy over the Tesseract for Android. It also reduces the processing time to less than half. All the preprocessing techniques were not added to the server, due to the late implementation. After the server was developed, the algorithms implemented were tested with the server, where the accuracy obtained with the preprocessing techniques improved to 40%.

In order to achieve even better results, the four frames prior of the actual photo, were included in the processing. These four frames suffers different processing operations, leading to an increase

of the accuracy to 55% and a decrease to 8% of 0 matches. An improvement of almost the double of the initial accuracy and a decrease of one third of the labels that return 0 matches. As this solution also increases the results obtained by the Android application, the OCR was kept in the Android so the user can use it even without an Internet connection.

## 6.2 Conclusions

The alarming explosion of diet related diseases is one of the reasons for the increasing concern from the population to improve their diet. Although almost all the products in the market contain nutrition labels, presenting information about the nutrients, the consumers show some problems when analyzing the data presented.

The main goal of this project was to develop a tool that helps the user making healthier choices, by means of an Android application. This application extracts automatically information from a nutritional label and presents that information in a cross-sectional shape, adding some assistance regarding the user's problems and needs in that task. The application also follows the new regulation regarding the presentation of the nutritional information, that most nutrition labels in the time of the development of this project didn't follow. Another goal of this project was to allow the user to compare between two products of the same category. Without any guidance, it is required nutritional knowledge and consideration about the role of each product and the overall values of the nutrients, for the user to be able to make this decision. In order to aid the user to make his decision, the application compares the selected products and displays the results in a form of a table containing the nutrient values. The application also displays a message with the healthier product.

To test the image processing operations and its impact on the accuracy of the OCR, fourteen images of five different smartphones were used. Unlike the study previously analysed of the Tesseract, the results obtained of the OCR were beneath *high accuracy* as it was advertised, with an accuracy of only 20%. This value is far from the high accuracy claimed by Google, but it is also the only OCR found that was free and open source. In order for the application retrieve better results a server was added to the system, showing a quick improvement on the OCR. After all the tests performed, the best results were found sending multiple frames of the same nutrition declaration and perform different algorithms in each one.

In the end, the accuracy was 55%, an increase of 83% considering the OCR without any preprocessing, and the time consumed can reach 10 seconds, which is longer than intended. Nevertheless, only 8% of the nutritional declarations tested retrieved 0 matches.

## 6.3 Future work

As opposed of the initial planning, the functionality of assessing the suitability of a given product, according to the personal data of the user, was not implemented in the application. This feature

## Discussion, conclusions and future work

can include treatment of specific diseases, diet restrictions, among other possibilities. So in the near future, this is a feature to be added in the application.

Another possible improvement for the near future, is to save the product information in a database, with the objective to check the nutritional information of a previously processed product or even use it to the comparison.

To increase the quantity of the products that can be compared, the product categories used in the application should be extended.

To improve the accuracy, some improvements or even new approaches to the image processing algorithms can be made. These operations also can be performed in the server instead of in the application, reducing drastically the time of the processing.

To improve the results of the Tesseract, this engine can be trained, but Google only advises the training when it uses a very unusual font or a new language, which it is not usual.

Another option is to use or develop another OCR engine, but at the time of the development of this project wasn't found any OCR engine free and open source. In order to have a better overall understanding of the role of the image resolutions and the quality of the camera with the results of the OCR, the set of images to test should be enlarged.

## Discussion, conclusions and future work

# References

- [AAAA14] Khalaf S Alkhalaf, Abdulelah I Almishal, Anas O Almahmoud, and Majed S Alotaibi. Ocr-based electronic documentation management system. 2014.
- [Aga06] Gady Agam. Introduction to programming with opencv. *Online Document*, 27, 2006.
- [AGH06] Alfonso Abizaid, Qian Gao, and Tamas L Horvath. Thoughts for food: brain mechanisms and peripheral energy balance. *Neuron*, 51(6):691–702, 2006.
- [Ber96] Kent C Berridge. Food reward: brain substrates of wanting and liking. *Neuroscience & Biobehavioral Reviews*, 20(1):1–25, 1996.
- [BGR07] Wojciech Bieniecki, Szymon Grabowski, and Wojciech Rozenberg. Image preprocessing for improving ocr accuracy. In *Perspective Technologies and Methods in MEMS Design, 2007. MEMSTECH 2007. International Conference on*, pages 75–80. IEEE, 2007.
- [Bir98] Leann L Birch. Development of food acceptance patterns in the first years of life. *Proceedings of the Nutrition Society*, 57(04):617–624, 1998.
- [BRHT04] N Bin Rais, M Shehzad Hanif, and Imtiaz A Taj. Adaptive thresholding technique for document image analysis. In *Multitopic Conference, 2004. Proceedings of INMIC 2004. 8th International*, pages 61–66. IEEE, 2004.
- [BW04] Swinburn Ba and James WPT. Diet, nutrition and the prevention of excess weight gain and obesity. *Public health nutrition*, 7(1a):123–146, 2004.
- [CCI13] Sílvia Coelho, Patrícia Carrilho, and Luís Inchaustegui. Management of hyperglycaemia in type 2 diabetic patients with chronic kidney disease. *Portuguese Journal of Nephrology & Hypertension*, 27:91–100, 2013.
- [CS05] Gill Cowburn and Lynn Stockley. Consumer understanding and use of nutrition labelling: a systematic review. *Public health nutrition*, 8(01):21–28, 2005.
- [CSSM11] Monica Caldeira, Miguel Sottomayor, and Diogo M Souza Monteiro. Portuguese retailers’ motivations to adopt front of pack nutrition labels: A qualitative analysis. In *2011 Annual Meeting, July 24-26, 2011, Pittsburgh, Pennsylvania*. Agricultural and Applied Economics Association, 2011.
- [CSV<sup>+</sup>15] Michelle Crino, Gary Sacks, Stefanie Vandevijvere, Boyd Swinburn, and Bruce Neal. The influence on population weight gain and obesity of the macronutrient composition and energy density of the food supply. *Current Obesity Reports*, pages 1–10, 2015.

## REFERENCES

- [DG01] Lijun Ding and Ardeshir Goshtasby. On the canny edge detector. *Pattern Recognition*, 34(3):721–725, 2001.
- [DTJT96] Øivind Due Trier, Anil K Jain, and Torfinn Taxt. Feature extraction methods for character recognition—a survey. *Pattern recognition*, 29(4):641–662, 1996.
- [eMa14] eMarketer. Worldwide smartphone usage to grow 25% in 2014. 2014.
- [EU11] EU. Regulamento (ue) n.º 1169/2011 do parlamento europeu e do conselho de 25 de outubro de 2011 relativo à prestação de informação aos consumidores sobre os géneros alimentícios. *Jornal Oficial da União Europeia*, 22:18–63, 2011.
- [FDA15] FDA. Factsheet on the new proposed nutrition facts label. Available at: <http://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/LabelingNutrition/ucm387533.htm>, note = Access date: 2015-02-11, 2015.
- [FDF14] FDF. Science behind guideline daily amounts. 2014.
- [FF15] Food and Drink Federation. Using the front of the pack label. Available at: [http://www.foodlabel.org.uk/label/gda\\_values.aspx](http://www.foodlabel.org.uk/label/gda_values.aspx), note = Access date: 2015-02-05, 2015.
- [FSS13] Atena Farahmand, Abdolhossein Sarrafzadeh, and Jamshid Shanbehzadeh. Document image noises and removal methods. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 1, 2013.
- [gBCL<sup>+</sup>10] S Storcksdieck genannt Bonsmann, L Fernández Celemín, Ainhoa Larrañaga, Salome Egger, Josephine M Wills, Charo Hodgkins, and MM Raats. Penetration of nutrition information on food labels across the eu-27 plus turkey. *European journal of clinical nutrition*, 64(12):1379–1385, 2010.
- [GFCW<sup>+</sup>10] Klaus G Grunert, Laura Fernández-Celemín, Josephine M Wills, Stefan Storcksdieck genannt Bonsmann, and Liliya Nureeva. Use and understanding of nutrition information on food labels in six european countries. *Journal of Public Health*, 18(3):261–277, 2010.
- [GJG07] Maya R Gupta, Nathaniel P Jacobson, and Eric K Garcia. Ocr binarization and image pre-processing for searching historical documents. *Pattern Recognition*, 40(2):389–397, 2007.
- [GKM13] A Geetha Kiran and S Murali. Automatic rectification of perspective distortion from a single image using plane homography. 2013.
- [GR03] Yolanda Gonzalez and Jairo Rocha. Some practical assignments in computer vision. In *Proceedings of the Workshop on Education and Practice in Artificial Vision*, pages 101–104. Citeseer, 2003.
- [GW07] Klaus G Grunert and Josephine M Wills. A review of european research on consumer response to nutrition information on food labels. *Journal of Public Health*, 15(5):385–399, 2007.

## REFERENCES

- [GZMS05] Grazielle Guimaraes Granada, Rui Carlos Zambiasi, Carla Rosane Barboza Mendonça, and Edilson Silva. Caracterização física, química, microbiológica e sensorial de geléias light de abacaxi. *Food Science and Technology (Campinas)*, 25:629–635, 2005.
- [Hea13] Department of Heath. Guide to creating a front of pack (fop) nutrition label for pre-packed products sold through retail outlets, 2013.
- [HEL<sup>+</sup>04] CR Hankey, S Eley, WS Leslie, CM Hunter, and MEJ Lean. Eating habits, beliefs, attitudes and knowledge among health professionals regarding the links between obesity, nutrition and health. *Public health nutrition*, 7(02):337–343, 2004.
- [HW12] S Hieke and JM Wills. Nutrition labelling. is it effective in encouraging healthy eating. *CAB Reviews*, 7(31):1–7, 2012.
- [HWA<sup>+</sup>13] James C Hersey, Kelly C Wohlgenant, Joanne E Arsenault, Katherine M Kosa, and Mary K Muth. Effects of front of package and shelf nutrition labeling systems on consumers. *Nutrition reviews*, 71(1):1–14, 2013.
- [JJ05] L Jagannathan and CV Jawahar. Perspective correction methods for camera based document analysis. In *Proc. First Int. Workshop on Camera-based Document Analysis and Recognition*, pages 148–154, 2005.
- [JSVR05] Charles Jacobs, Patrice Y Simard, Paul Viola, and James Rinker. Text recognition of low-resolution document images. In *Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on*, pages 695–699. IEEE, 2005.
- [KLM09] Andrew Kae and Erik Learned-Miller. Learning on the fly: font-free approaches to difficult ocr problems. In *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on*, pages 571–575. IEEE, 2009.
- [LDD05] Jian Liang, Daniel DeMenthon, and David Doermann. Flattening curved documents in images. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 338–345. IEEE, 2005.
- [Lee03] Jeong Sook Lee. The effects of gender, obesity rate, nutrition knowledge and dietary attitude on the dietary self-efficacy of adolescents. *Korean Journal of Community Nutrition*, 8(5):652–657, 2003.
- [LLKK05] Chulhan Lee, Sanghoon Lee, Jaihie Kim, and Sung-Jae Kim. *Preprocessing of a fingerprint image captured with a mobile camera*, pages 348–355. Springer, 2005.
- [LS99] GS Lehal and Chandan Singh. Feature extraction and classification for ocr of gurmukhi script. *VIVEK-BOMBAY*, 12(2):2–12, 1999.
- [LSAR00] Tobias Leidinger, Lübmira Spassova, Andreas Arens, and Norbert Rösch. A mobile user interface for semi-automatic extraction of food product ingredient lists. 2000.
- [Mah13] Lisa Mahapatra. Android vs. ios: What's the most popular mobile operating system in your country? *International Business Times*, 2013.
- [Man10] Bidisha Mandal. Use of food labels as a weight loss behavior. *Journal of Consumer Affairs*, 44(3):516–527, 2010.

## REFERENCES

- [McD13] John McDermott. A majority of u.s. mobile users are now smartphone users. *Advertising Age*, 2013.
- [MHVCV10] Anke Möser, Christine Hoefkens, John Van Camp, and Wim Verbeke. Simplified nutrient labelling: consumers’ perceptions in germany and belgium. *Journal für Verbraucherschutz und Lebensmittelsicherheit*, 5(2):169–180, 2010.
- [Nes13] Marion Nestle. *Food politics: How the food industry influences nutrition and health*, volume 3. Univ of California Press, 2013.
- [NESM01] Premkumar Natarajan, Baback Elmieh, Richard Schwartz, and John Makhoul. Videotext ocr using hidden markov models. In *Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on*, pages 947–951. IEEE, 2001.
- [NKP99] Marian L Neuhouser, Alan R Kristal, and Ruth E Patterson. Use of food nutrition labels is associated with lower fat intake. *Journal of the American Dietetic Association*, 99(1):45–53, 1999.
- [Nut11] NutriScanner. Nutriscanner nutrition facts label ocr scanner. Available at: [http://download.cnet.com/NutriScanner-Nutrition-Facts-Label-OCR-Scanner/3000-2129\\_4-75568005.html](http://download.cnet.com/NutriScanner-Nutrition-Facts-Label-OCR-Scanner/3000-2129_4-75568005.html), note = Access date: 2015-02-05, 2011.
- [Onl10] The Star Online. Nutrition information. Available at: <http://www.thestar.com.my/Story/?file=/2010/10/24/health/7251290>, note = Access date: 2015-03-01, 2010.
- [Ope15] OpenCV. Image thresholding. Available at: [http://docs.opencv.org/master/d7/d4d/tutorial\\_py\\_thresholding.html](http://docs.opencv.org/master/d7/d4d/tutorial_py_thresholding.html), note = Access date: 2015-02-05, 2015.
- [PCAAL00] JAJC Perez-Cortes, J Amengual, Joaquim Arlandis, and Rafael Llobet. Stochastic error-correcting parsing for ocr post-processing. In *Pattern Recognition, International Conference on*, volume 4, pages 4405–4405. IEEE Computer Society, 2000.
- [PPP12] Chirag Patel, Atul Patel, and Dharmendra Patel. Optical character recognition by open source ocr tool tesseract: A case study. *International Journal of Computer Applications*, 55(10):50–56, 2012.
- [Ras55] Nicolas Rashevsky. Life, information theory, and topology. *The bulletin of mathematical biophysics*, 17(3):229–235, 1955.
- [RFP13] Sónia Ribeiro, Cláudia Furtado, and João Pereira. Association between cardiovascular disease and socioeconomic level in portugal. *Revista Portuguesa de Cardiologia (English Edition)*, 32(11):847–854, 2013.
- [RMD<sup>+</sup>09] Ronielli Cardoso REIS, Valéria Paula Rodrigues MINIM, Beatriz Rodrigues Pimentel DIAS, José Benício Paes CHAVES, and Luis Antonio MINIM. Impacto da utilização de diferentes edulcorantes na aceitabilidade de iogurte “light” sabor morango. *Alimentos e Nutrição Araraquara*, 20(1):53–60, 2009.

## REFERENCES

- [Rui97] Arnout C Ruifrok. Quantification of immunohistochemical staining by color translation and automated thresholding. *Analytical and quantitative cytology and histology/the International Academy of Cytology [and] American Society of Cytology*, 19(2):107–113, 1997.
- [Sau13] Portal da Saude. Plano nacional de saude 2011-2016. 2013.
- [SD01] Mauritius Seeger and Christopher Dance. Binarising camera images for ocr. In *Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on*, pages 54–58. IEEE, 2001.
- [Smi07a] Ray Smith. An overview of the tesseract ocr engine. In *ICDAR*, volume 7, pages 629–633, 2007.
- [Smi07b] Ray Smith. Tesseract ocr engine. *Lecture. Google Code. Google Inc*, 2007.
- [Sta15] StatCounter. Globalstats, 2015.
- [TMBY13] Sandeep Tiwari, Shivangi Mishra, Priyank Bhatia, and Praveen Km Yadav. Optical character recognition using matlab. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, 2(5):579–582, 2013.
- [WHO00] WHO. *Obesity: preventing and managing the global epidemic*. World Health Organization, 2000.
- [WHO09] WHO. 2008-2013 action plan for the global strategy for the prevention and control of noncommunicable diseases: prevent and control cardiovascular diseases, cancers, chronic respiratory diseases and diabetes. 2009.
- [WHO15] WHO. Healthy diet. Available at: <http://www.who.int/mediacentre/factsheets/fs394/en/>, note = Access date: 2015-02-11, 2015.
- [Wil99] Melvin H Williams. *Nutrition for health, fitness and sport*. WCB/McGraw-Hill, 1999.
- [YY11] Linda He Yi and Feiqiao Brian Yu. Automatic grocery shopping assistant. 2011.

## REFERENCES

# Appendix A

## Appendix 1

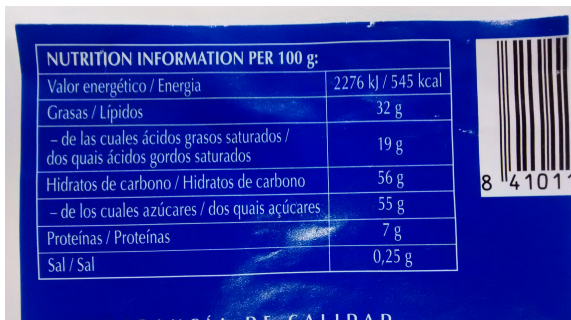
### A.1 Set of Images used in the tests

This is the set corresponding to MotoG.



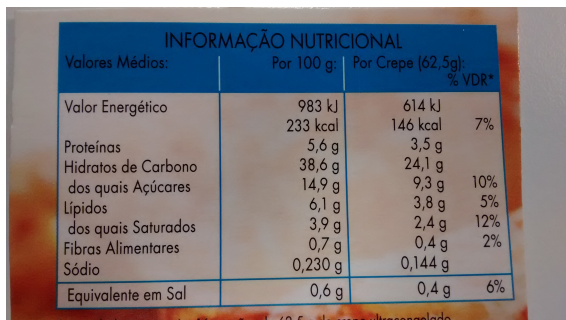
INFORMAÇÃO NUTRICIONAL	Por 100 g	Por 23,5g
Valor energético	1576 kJ 373 kcal	370 kJ 88 kcal
Lípidos dos quais ácidos gordos saturados	7,0 g 3,3 g	1,6 g 0,8 g
Hidratos de carbono dos quais açúcares	69,7 g 24,6 g	16,4 g 5,8 g
Fibra	5,0 g	1,2 g
Proteínas	5,4 g	1,3 g
Sal	0,95 g	0,22 g

Figure A.1: cookies1



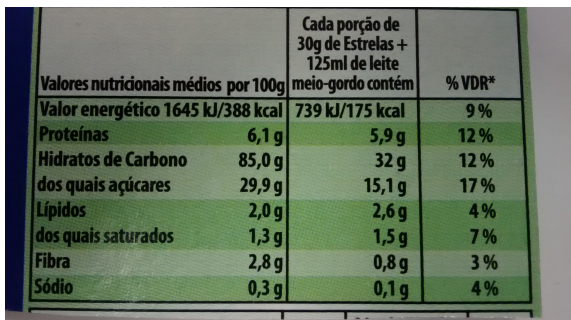
NUTRITION INFORMATION PER 100 g:	
Valor energético / Energia	2276 kJ / 545 kcal
Grasas / Lípidos	32 g
- de las cuales ácidos grasos saturados / dos quais ácidos gordos saturados	19 g
Hidratos de carbono / Hidratos de carbono	56 g
- de los cuales azúcares / dos quais açúcares	55 g
Proteínas / Proteínas	7 g
Sal / Sal	0,25 g

Figure A.2: cookies2



INFORMAÇÃO NUTRICIONAL			
Valores Médios:	Por 100 g:	Por Crepe (62,5g):	% VDR*
Valor Energético	983 kJ 233 kcal	614 kJ 146 kcal	7%
Proteínas	5,6 g	3,5 g	
Hidratos de Carbono dos quais Açúcares	38,6 g 14,9 g	24,1 g 9,3 g	10%
Lípidos dos quais Saturados	6,1 g 3,9 g	3,8 g 2,4 g	5% 12%
Fibras Alimentares	0,7 g	0,4 g	2%
Sódio	0,230 g	0,144 g	
Equivalente em Sal	0,6 g	0,4 g	6%

Figure A.3: cookies3



Valores nutricionais médios por 100g	Cada porção de 30g de Estrelas + 125ml de leite meio-gordo contém	% VDR*
Valor energético 1645 kJ/388 kcal	739 kJ/175 kcal	9%
Proteínas 6,1 g	5,9 g	12%
Hidratos de Carbono 85,0 g	32 g	12%
dos quais açúcares 29,9 g	15,1 g	17%
Lípidos 2,0 g	2,6 g	4%
dos quais saturados 1,3 g	1,5 g	7%
Fibra 2,8 g	0,8 g	3%
Sódio 0,3 g	0,1 g	4%

Figure A.4: cookies4

Appendix 1

cálcio), sal, levedantes (bicarbonato de sódio, bicarbonato de amônio), soro de **leite** em pó, canela em pó 0,03 %, emulsionante (lecitina de girassol).  
**Pode conter ovo.**

DECLARAÇÃO NUTRICIONAL			
	100 g	2 bolachas (24,6 g)	* % / 2 bolachas
Energia	1620 kJ 385 kcal	398 kJ 95 kcal	5 %
Lípidos	8,5 g	2,1 g	3 %
dos quais ácidos gordos saturados	3,4 g	0,8 g	4 %
Hidratos de carbono	73,5 g	18 g	7 %
dos quais açúcares	25,5 g	6,3 g	7 %
Fibra	3,3 g	0,8 g	-
Proteínas	4,6 g	1,1 g	2 %
Sal	0,60 g	0,15 g	3 %

\* Doses de referência para um adulto médio (8400 kJ / 2000 kcal).  
**12 bolachas / embalagem**

Figure A.5: cookies5

Informação Nutricional/Nutritional Information	Por 100g
Valor energético/Energy	423 Kcal 1760 KJ
Lípidos/Fat	9 g
dos quais saturados/ Of which saturated	4.08 g
Hidratos de carbono/Carbohydrate	79.9 g
dos quais açúcares/Of which sugars	28.6 g
Proteínas/Protein	7.6 g
Sódio/Sodium	0.26 g

Conservar em lugar fresco e seco ao abrigo da luz

Figure A.6: cookies6

Declaração Nutricional Nutrition Declaration	por 100g per 100g	por 1 bolacha (aprox. 3,1g) per 1 biscuit (approx. 3,1g)	DR* DR*
Energia / Energy	1743kJ/414kcal	54kJ/13kcal (0,6% DR*)	8400kJ/ 2000kcal
Lípidos / Lipids	10,0g	0,31g (0,4% DR*)	70 g
Dos quais / From which: - Ácidos gordos saturados/ Saturated fatty acids	4,9g	0,15g (0,8% DR*)	20 g
Hidratos de Carbono / Carbohydrates	66,9g	2,07g (0,8% DR*)	260 g
Dos quais / From which: - Açúcares / Sugars	6,9g	0,21g (0,2% DR*)	90 g
Fibras	5,6	0,17g	
Proteínas / Proteins	11,2g	0,35g (0,7% DR*)	50 g
Sal / Salt	0,6g	0,02g (0,3% DR*)	6 g

DR\* - Doses de referência para um adulto médio (8400 kJ / 2000 kcal)

Figure A.7: cookies7

DECLARAÇÃO NUTRICIONAL			
	Por 100 g	Por 3 bolachas (30 g)	%* Por 3 Bolachas (30 g)
Energia	1765 kJ 420 kcal	530 kJ 126 kcal	6 %
Lípidos	12 g	3,6 g	5 %
dos quais ácidos gordos saturados	2,9 g	0,9 g	5 %
Hidratos de Carbono	62 g	18,5 g	7 %
dos quais açúcares	15,5 g	4,7 g	5 %
Fibra	8,6 g	2,6 g	-
Proteínas	11,5 g	3,5 g	7 %
Sal	0,53 g	0,16 g	3 %
Cálcio	320 mg	39 % VRN**	95 mg 12 % VRN**
Ferro	4,4 mg	32 % VRN**	1,3 mg 9 % VRN**

Figure A.8: cookies8

Appendix 1

Valores nutricionais médios	ø/100g	ø/Porção 30g (+125ml leite 1/2 gordo)	%RI
Energia	1598 kJ/377 kcal	720 kJ/170 kcal	9%
Lípidos	1,1g	2,3g	3%
dos quais saturados	0,5g	1,3g	7%
Hidratos de carbono	84,0g	31g	12%
dos quais açúcares	7,1g	8g	9%
Fibra	3,4g	1,0g	--
Proteínas	6,0g	5,9g	12%
Sal	1,80g	0,53g	9%

RI (reference intake) = Dose de Referência para um adulto médio (8400 kJ/2000 kcal) (DR)

Figure A.9: cookies9

leite pasteurizado desnatado, leite, fermentos lácteos, sal, açúcar, amido de milho, aspartame e acesulfame K. Contêm uma fonte de fenilalanina. Isento de glúten.

Valores médios	por 100ml	por unid.175ml
energia (kJ/kcal)	121/29	212/50 3%*
lipídios (g)	0,2	0,4 1%*
-- dos quais: ácidos gordos saturados (g)	0,1	0,2 1%*
hidratos de carbono (g)	3,3	5,8 2%*
-- dos quais: açúcares *** (g)	3,3	5,8 6%*
proteínas (g)	3,4	6,0 12%*
sal (g) (1)	0,13	0,22 4%*
vitaminas e sais minerais		
-- riboflavina (mg)	0,25 18%*	0,44 31%*
-- vitamina B12 (µg)	0,23 9%*	0,40 16%*
-- cálcio (mg)	118 15%*	207 26%*
-- fósforo (mg)	92,0 13%*	161 23%*
-- iodo (µg)	28,0 19%*	49,0 33%*

Valor de Referência do Nutriente (VRN) \*\* da Dose de Referência (DR) para um adulto médio (8400kJ/2000kcal). \*\*\* Contêm açúcares naturalmente presente

Figure A.10: yoghurt1

Valores médios	por 100ml	por unid.175ml
energia (kJ/kcal)	121/29	212/50 3%*
lipídios (g)	0,2	0,4 1%*
-- dos quais: ácidos gordos saturados (g)	0,1	0,2 1%*
hidratos de carbono (g)	3,3	5,8 2%*
-- dos quais: açúcares *** (g)	3,3	5,8 6%*
proteínas (g)	3,4	6,0 12%*
sal (g) (1)	0,13	0,22 4%*
vitaminas e sais minerais		
-- riboflavina (mg)	0,25 18%*	0,44 31%*
-- vitamina B12 (µg)	0,23 9%*	0,40 16%*
-- cálcio (mg)	118 15%*	207 26%*
-- fósforo (mg)	92,0 13%*	161 23%*
-- iodo (µg)	28,0 19%*	49,0 33%*

Figure A.11: yoghurt2

Declaração Nutricional (valores médios)	por 100g	por UNIDADE	%DR* por UNIDADE
Energia	652kJ/156kcal	749kJ/180kcal	9%
Lípidos	8,9g	10,3g	15%
dos quais ácidos gordos saturados	5,8g	6,6g	33%
Hidratos de Carbono	15,4g	17,8g	7%
dos quais açúcares	14g	16,1g	18%
Fibra	0,2g	0,2g	-
Proteínas	3g	3,4g	7%
Sal	0,07g	0,08g	1%

\*DR - Doses de Referência para um adulto médio (8400kJ/2000kcal).

Figure A.12: yoghurt3

Información nutricional (ES)	Valores nutricionales médios (PT)	ø/100g	ø/125g	% GDA*
Valor energético	Valor energético	407 kJ/96 kcal	509 kJ/120 kcal	6%
Proteínas	Proteínas	4,1g	5,1g	10%
Hidratos de carbono de los cuales, azúcares	Hidratos de carbono dos quais açúcares	16,0g/15,0g	20,0g/18,8g	7%/21%
Grasas de las cuales, saturadas	Lípidos dos quais saturados	1,4g/1,0g	1,8g/1,3g	3%/7%
Fibra alimentaria	Fibras alimentares	0g	0g	0%
Sódio	Sódio	0,06g	0,08g	3%

\* ES Cantidad diaria orientativa para adultos  
PT Valores diarios de referencia para un adulto  
Más información en: [www.lidl.es](http://www.lidl.es)

Figure A.13: yoghurt4

Valores médios	por 100g	por unid.125g
energia (kJ/kcal)	160/38	201/47 2%*
lipídios (g)	0,2	0,3 0%*
dos quais: ácidos gordos saturados (g)	0,1	0,2 1%*
hidratos de carbono (g)	4,6	5,8 2%*
dos quais: açúcares *** (g)	4,6	5,8 6%*
proteínas (g)	4,4	5,5 11%*
sal (g) (1)	0,15	0,19 3%*
vitaminas e sais minerais		
-- riboflavina (mg)	0,30 21%*	0,38 27%*
-- vitamina B12 (µg)	0,47 19%*	0,59 24%*
-- cálcio (mg)	163 20%*	204 25%*
-- fósforo (mg)	128 18%*	160 23%*
-- iodo (µg)	21,9 15%*	27,4 18%*

Figure A.14: yoghurt5