An Embedded Software Development Package Platform for Cloud OCR

Dan Marius Dobrea[1], Paul Boț[2]
Electronics, Telecommunications and Information Technology, "Gheorghe Asachi" Technical University, Iași, Romania
[1]mdobrea@etti.tuiasi.ro, [2] botz_paul@yahoo.com

***Abstract*—In a world focused on the need of selling an increasingly large quantity of goods, quality, visual appearance, texture, flavor, taste etc. become central elements that support the global economic development. As a result, different types of food additives were and are still designed and used to fulfill such objectives. In what follows, the aim of this paper is to present the design and implementation of a new software package for embedded systems (smart phones) able to support a customer in the process of food purchasing. The developed software package runs on an Android platform and uses a cloud optical character recognition (OCR) algorithm. The information related to the product of interest and supplied at request to the customers are: the name and the code of the product's additives, the risks associated with each food additive, the source of each food components, the limit of its daily consumption and the side effects, diet restrictions etc.**

I. Introduction

Food additives are different types of chemical substances used during the entire flow of production, processing, packaging and storage of the food. The food additives (**FA**) play different roles; e.g. they are used in food (a) to preserve or enhance flavor, taste, texture or appearance, (b) to enhance the nutrient value of our food (e.g. different types of vitamins and minerals), (c) to make possible for food to be fresh and healthy, (d) to make food more appealing etc.

To each FA a unique number is assigned. In both, European Union (EU) and Switzerland, all approved FA have a specific number prefixed by "E" – where "E" stands for Europe. Also, the E numbers are used in food labelling in other zones and countries, including Australia, Israel, New Zealand, South Africa and the countries from Cooperation Council for the Arab States of the Gulf. In some other countries, only a number is used. The database for FA numbering is now adopted and completed by the Codex Alimentarius in order to worldwide identify all types of FA, regardless of whether they are approved for use or not.

Safety assessment and approval of FA are the responsibility of both the European Food Safety Authority (EFSA) in EU and the Food and Drug Administration (FDA) in the USA. The regulations of these organizations require that all food ingredients (including FA) be listed on the package label.

Today, in the entire world a significant controversy was and continues to be associated with the risks and the benefits of different types of FA. Even if there are a large number of FA that are considered "good" and approved in the EU and USA, there are also some of them that are in dispute or approved either in EU only or exclusively in USA.

FDA allowed several FA that are banned in many other developed countries, like: E443 (brominated vegetable oil – an emulsifier not approved in EU, UK and Japan), different types of improved agents like E924 and E927a (not approved in EU), colorants like E110 etc. Even in the USA, due to the concerns that some specific FA can cause harmful effects, several companies remove them from their products. The negative influence of the brominated vegetable oil (E443) in the development of birds and animals [1] and the toxicological concerns regarding the accumulation of bromine in fatty tissues both in experimental animals and isolated human cases [2], have determined PepsiCo™ (since May 2014) and Coca-Cola™ (since the end of 2014) to remove it from the most of the drinks. The International Agency for Research on Cancer considers E924 as a carcinogenic substance to humans – classified as a category 2B carcinogen [3].

Even in Europe, there are divergences in the legislations of its different countries. Thus, for example, while E110 is approved in the EU it is banned in Austria and Norway (a country that adopted most of the EU legislation due to both its membership in the European Economic Area and its affiliation in the European Free Trade Association).

Other FA, currently approved for use in Australia, Canada, EU, New Zealand and USA (e.g., artificial colors or a sodium benzoate preservative – or both), present the evidence to produce an increased hyperactivity in the 3-year-old and 8/9-year-old children from the all general population [4].

Even if most of the additives in food are harmless, there are a number of FA that can generate a cumulative effect in time and, thus, cause different severe illnesses, like cancer.

The first goal of our research, presented here, is to develop a system able to support a customer when he/she acquires different food products. For this, the system has to be portable, able to "read" the label from the package, "understand" the meaning of different components and support the customer with information regarding: the risks associated with each FA, limit of daily consumption, side effects, diet restrictions etc.

The second goal of the paper consists in proposing and implementing the previous system in such a specific way as an independent development platform. This platform will have to sustain different types of OCR applications and algorithms on mobile embedded systems like smart phones. The requirements for this platform imply abilities to implement and test various types of OCR algorithms quickly, in real environments, based on cloud support and in a real-time manner.

II. THE SYSTEM'S REQUIREMENTS AND THE PROPOSED SOLUTIONS

The system and the software package (**SP**) that should fulfill all the previous requirements must, also, be very easy to use and target a large class of consumers. For this reason, even from the beginning we chose a smartphone with Android operating system (OS). According to the International Data Corporation (IDC) [5], the Android OS has owned, in the third quarter of 2014, over 84% of the market share.

For the simplicity of the application we decided to use the phone's camera to acquire the image of the food label. Just by scanning the food package, the user can be informed about all the FA contained in the product as well as about the potential risks that its consumption implies.

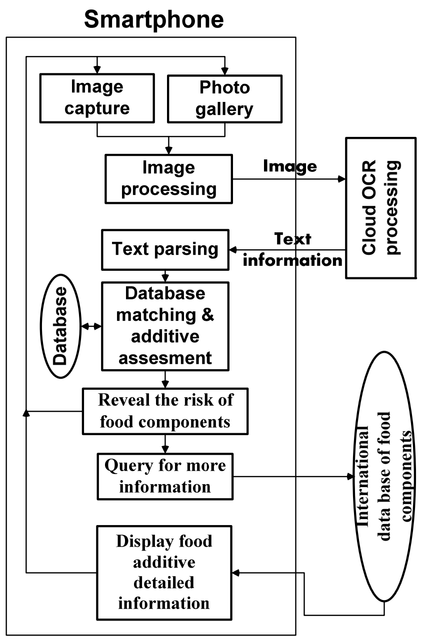

Figure 1. The software data flow and architecture

Another challenge of the project continues to be the huge quantity and the dynamics of the information as well as the very specific and technical characteristic of this information associated with the list of FA. Thus, for example, this list was enhanced recently by the Codex Alimentarius. Also, the E database is in a continuous process of adjustment, new compounds being inserted, while others are discovered as having a high risk factor to humans or even as being harmfully. All these imply, as a necessity, the use of an international database (**DB**) of food components that should be, in turn, supported by a prestigious team in the field of food science. We chose the Wageningen University's DB, from Netherlands, one of the leading universities in Food Science in the world.

Based on previous experience [6], we decided to develop the SP as an independent development platform, able to sustain the implementation and testing of different types of image processing algorithms and OCR methods quickly and in real-time. As a result, this research focused on the development of the frame system, presented in Fig. 1. Also, due to the computational load and the complexity of the algorithms we chose to move in cloud [7] the OCR process. Nevertheless, in order to prove the main concepts we had, also, to focus on developing and implementing a full functional prototype. This is the reason for which we chose to use, for beginning, a very strong character recognition engine, namely, a web service from ABBYY Software Company. This was mainly justified by the label printed on the food packages that is generally of a low quality, with frequently very small written text.

III. THE SOFTWARE TECHNICAL DESCRIPTION

A. *Software architecture*

In our application, the user has the following two options: 1) to take an image of a product label, using the camera from its mobile device, and 2) to load an already existing image from the gallery, see Fig. 1. The last option is especially useful when the user subscription does not include in his monthly price plan any data connection. In particular, the second method allows finding information for product ingredients by uploading the image when the user has access to a Wi-Fi network.

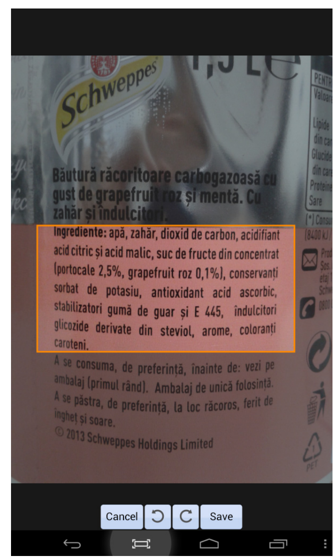

Figure 2. Selecting ingredients by cropping the image

In the software part the image processing step is performed in two phases. In a first step, the image acquisition module adjusts the brightness, saturation, contrast, hue, sharpness, white balance, exposure time and the color effect of the image. In the second stage, because the ingredients on the package can be formatted in different ways, the program must allow to select only the label's section in which the ingredients are listed. For this, an image cropping function was implemented, Fig. 2.

Using the OCR process, the text is obtained from the selected image. In a first step, the text obtained by the OCR module is divided in words and, then, a list of food ingredients is created. For this operation, the text parser module follows this rule: two distinct ingredients are divided either by a comma, a point or a semicolon; if more than one word occurs between two separator-marks these words will be consider together as composing an ingredient name. Next, each ingredient resulting from the processed text is searched in a local DB or on the web. To support the customer decision both methods are used. Based on the customer requirements, only the risk factor of each ingredients are presented (Fig. 3) or a detailed information is displayed (Fig. 4).

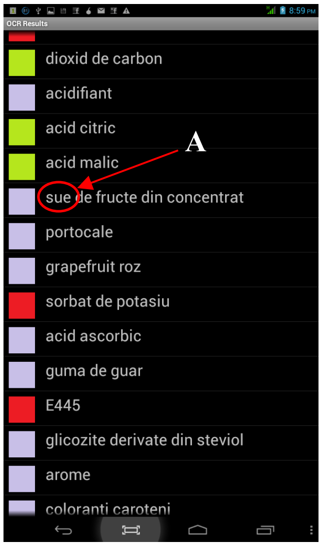

Figure 3. List of ingredients determined by cloud OCR process

B. *User interface*

The graphical user interface of this application is a minimal one. The main view that is the entry point into application contains only two buttons. The "window" concept from Windows OS world is named "view" in the Android world. These buttons allows the user to capture a photo or to select an existing image from the gallery. These actions are performed using the following two intents: (a). MediaStore.ACTION_IMAGE_CAPTURE and, respectively, (b). Intent.ACTION_PICK.

The information are displayed to the customer in two views: (a) a list of all ingredients with the degree of risk they pose to the health, Fig. 3 and (b) a detailed information regarding a single ingredient, namely, the ingredient selected by the user from the previous view.

C. *OCR cloud processing*

Currently, the OCR technology has evolved greatly and the state of the art algorithms can achieve recognition rates of over 98% - 99% for the characters in the Latin alphabet. We chose the ABBYY OCR web service. The advantages of using the ABBYY web services are: (a) the highly-accurate text recognition rate, supported by a powerful engine developed and improved over the last 20 years and (b) the abilities to recognize food labels printed in 190 languages. Also, the availability of the RESTful (**Re**presentational **S**tate **T**ransfer – REST) API provided by ABBYY removes all the compatibility issues that could arise with other platforms like iOS or Windows Phone. With this web service, the OCR process for any image is very simple and implies the usage of three methods: process image (***processImage***), get status task (***getStatusTask***) and download the results.

Figure 4. Detailed information about the selected ingredient

The image processing takes the following steps: loads an image into the cloud, creates a processing task and returns the ID of the task. With this ID the recognition progress can be assessed. This processing flow is called by accessing the resource ***processImage*** by POST method:

[POST] http(s)://cloud.ocrsdk.com/processImage

To this type of requests some other parameters could be provided such as: the image to be loaded, the language of the text or the type of the result – if the result should be returned in a particular format.

Based on the ID returned by the previous method, the ***getStatusTask*** method allows the software module to know the OCR process state, like:

[GET] http://cloud.ocrsdk.com/getTaskStatus?taskId=c3...def

Finally, the text resulting from the OCR process is achieved by a simple call that uses a GET method; this method makes use of a web address specified by a parameter returned by the previous function.

D. *Getting information on ingredients*

Once the ingredients have been obtained in the text format, the software can proceed to the next step, namely, the collection of information about each ingredient.

The risk factor of each of the ingredients is obtained by querying a local database (DB). If one ingredient is not found locally the search will be done on the web.

TABLE I. SAMPLE ENTRIES IN THE RISK DATABASE

ID	DENUMIRE	RISC
	…	
E102	tartrazina	2
E127	eritrozina	1
	…	

The local DB is of SQLite type and it contains a table with the following structure: ID CHAR(5) PRIMARY KEY, DENUMIRE CHAR (40), RISC INT. ID is the primary key of the table and it is given by the "E number" of the FA. DENUMIRE is the full name of the FA and RISC is a field of integer type that stores the FA risk factor. An example regarding the contents of this minimal DB is presented in Table 1. The risk factor can range from -1 to 3, as follows: (a) -1 - unknown risk, (b) 0 - not at risk for health, (c) 1 - medium risk: never exceed the recommended daily dose and (d) 2 - high risk: commonly consumed may have some side effects or even causes cancer. The -1 value is used each time when an ingredient is found neither in the local DB nor on the web DB. Since on the various labels the FA may be written either using the full name or the E number, the searching for the risk factor is done by comparing the search key to the ID field and, respectively, DENUMIRE field values. The local DB is also used to determine the E number associated with the entire name of an FA. This is useful in obtaining detailed information about each ingredient, as it is presented in Fig. 4. After having determined the risk factor for each ingredient from the list, this information is displayed graphical, as a square of specific color; this one is placed on the left side of each ingredient, see Fig. 3. The meaning of the colors is: gray - unknown risk, green - no risk, yellow - medium risk and red - high risk.

Whenever the user selects a specific ingredient, the software opens a new view with several detailed information about the ingredient, Fig. 4. This information is obtained from the Internet, using the web pages under the www.food-info.com domain. In this domain, the information regarding FA are stored in HTML documents as follows: *http://www.food-info.net/ro/e/e[E_number].htm*. E_number is the number for a specific FA. To extract information from the HTML documents we are using a library by which someone can easily parse a HTML document type: Jsoup.

IV. RESULTS

There are a large number of factors that can alter the OCR performance. In our particular case the main factors are skew degree and image: resolution, clarity and quality. Part of these factors (e.g. clarity and quality) is very hard to quantify because there are no good metrics to define them.

From our experience, a digital camera with at least 6-megapixel resolution is required. One central factor for clarity is given by the image stabilization technique; otherwise the user should take 2-3 shots of the food label in order to obtain an un-blurred image.

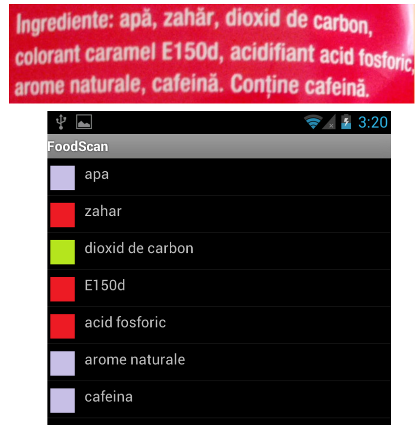

Figure 5. (a) The image and (b) the list of ingredients

In the case of a clear and a mean/high quality image, the entire proposed processing chain provides great results (see Fig. 2 and Fig. 3 or Fig. 5). The framework structure of the application works very well and it is able to support all these algorithms. The entire processing chain takes around 3 to 4 seconds in order to get the results.

On quality images, like the ones presented in Fig. 2 and Fig. 5(a), the mean correct recognition rates are 99.4% (computed from a set of 20 images). From time to time small errors happen – like the one presented in Fig. 3, A. The easiest solution to deal with this kind of errors is to develop a text parsing solution.

V. CONCLUSIONS

The technological development achieved in the last decade has led to an overwhelming progress for different types of software and hardware technologies. This revolution offers to all those who are interested, all the "bricks" that are necessary to develop professional and innovative applications such as the one presented in this paper.

REFERENCES

[1] I. D. Rae, “The Trouble with Bromine: Health and Environmental Impacts of Organobromine Compounds”, Global Environ., vol. 7, no. 1, pp. 106-133, 2014

[2] B. Z. Horowitz, “Bromism from excessive cola consumption”, J Toxicol Clin Toxicol., vol. 35, no. 3, pp. 315-20, 1997

[3] International Agency for Research on Cancer, Summaries & Evaluations: potassium bromate - group 2B, Vol. 73, 1999, pp. 481

[4] D. McCann, A. Barrett, A. Cooper, D. Crumpler, L. Dalen, K. Grimshaw, E. Kitchin, L. Lok, L. Porteous, E. Prince, E. Sonuga-Barke, J.O. Warner, and J. Stevenson, “Food additives and hyperactive behaviour in 3-year-old and 8/9-year-old children in the community: a randomised, double-blinded, placebo-controlled trial”, Lancet, vol. 370, no. 9598, pp. 1560-1567, 3 Nov. 2007

[5] IDC, Smartphone OS Market Share on Q3 of 2014, http://www.idc.com/prodserv/smartphone-os-market-share.jsp

[6] D. M. Dobrea, and M. C. Dobrea, “EEG classification system – from an universal system implementation to a particular signal modeling”, Proceedings of the Romanian Academy - Series A: Mathematics, Physics, Technical Sciences, Information Science, vol. 10, no. 2, pp. 197-204, May–August 2009

[7] D. M. Dobrea, D. Maxim, and Ş. Ceparu, “A face recognition system based on a Kinect sensor and Windows Azure cloud technology”, International Symposium on Signals, Circuits and Systems, July 11-12, 2013, Iași, România