

Proposal and Evaluation of Optical Sensor to Identify Liquids in Liquid Intake Detection System Using Smart Bottles

Sandra Viciano-Tudela¹, Paula Navarro-Garcia¹, Lorena Parra¹, Sandra Sendra¹, Jaime Lloret^{1*}

¹Instituto de Investigación para la Gestión Integrada de Zonas Costeras, Universitat Politècnica de València, 46730 Grau de Gandia, Spain
Email: svictud@upv.es, paunagar@alumni.upv.es, loparbo@doctor.upv.es, sansenco@upv.es, jlloret@dcom.upv.es

Abstract— Dehydration poses health risks, leading to confusion, falls, and even death. Tracking water intake, especially among at-risk groups, is vital. Smart bottles with sensors offer a solution for estimating and monitoring fluid consumption efficiently and widely. The paper aims to enhance the existing liquid intake detection systems by developing an optical sensing element to differentiate various liquids for precise hydration assessment. The study evaluates data processing techniques, including classic statistics, Discriminant Analysis, and Artificial Neural Networks, to classify liquids. The system is based on an ESP32 node integrated into the smart bottle as an Internet of Things device with communication capabilities with other wearable devices. A total of 7 different liquids are included in the conducted experiments. The data-gathering process is repeated several times to generate training and verification datasets. The results indicate that it is possible to differentiate the liquids using a reduced number of light wavelengths, white and purple. All analyzed techniques offered good results. Discriminant Analysis is the most effective classification approach with 100% accuracy. Nevertheless, if distinguishing between different types of teas is not necessary, thresholds based on statistical tools can be employed using fewer computation resources.

Keywords-RGB sensor; photoreceptors; Discriminant Analysis; Artificial Neural Network; Internet of Things

I. INTRODUCTION

Dehydration is a very important health problem, which can cause: confusion, falls, hospitalization, and, in the worst cases, death. That is why it is important to keep track of daily water consumption, especially in risk groups, such as the elderly and people with diseases that affect fluid regulation [1]. Despite the fact that there are groups that have priority, citizens' awareness of water consumption has increased since it has been shown that staying well-hydrated is essential both for our physical well-being and for cognitive health [2]. A survey was carried out where it was observed that 70% of those surveyed under the age of 50 admitted they had forgotten to drink water or feared they had not done so due to their hectic lifestyle [3]. Being the consumption of water an issue that affects not only risk groups but also the rest of the population.

Regarding the current solutions for stimulating fluid intake, multiple proposals can be found based on technological systems. On the one hand, there are systems, mainly integrated into smart devices, such as smartphones or smartwatches, that indicate the necessity of drinking every certain given period of time [4]. On the one hand, there are

systems that monitor drinking habits based on sensors embedded in wearable devices. Nevertheless, in this case, most of these systems are designed to identify problems linked with alcohol consumption [5]. The use of sensors embedded in a smart bottle that determines fluid intake is limited to medical surveillance [6], athletes [7], or elderly people [8]. For all this, it would be convenient to find a system that allows estimating daily water intake efficiently and simply with applications from hospitals and other health centres to the rest of the population.

Smart bottles are portable devices capable of detecting the type and amount of liquid you ingest, thanks to the use of sensors and physiological parameters [2]. Users can use and manage them in different ways: using a smartphone or uploading the data to the cloud [9]. Regarding the technologies used, the Liquid Intake Detection System (LIDS) provides real-time monitoring of the type and volume of fluid intake. For this, the system designs a detection module comprising ultrasonic, Red-Green-Blue (RGB) colour, temperature and accelerometer sensors, as well as a computational framework for classifying the type of fluid intake [2].

The aim of this paper is to develop the RGB sensing element of the LIDS, which is capable of distinguishing between different liquids. Among existing systems, smart bottles are focused on quantifying the liquids without distinguishing the type of liquid. Determining the sort of liquid is important in order to establish hydration needs. Therefore, the proposed system is tested with 7 different sorts of liquids in this paper. As the optical sensor, an RGB Light Emitting Diode (LED) module is used and configured to emit 7 different light colours and a photodetector. Calibration and verification tests are conducted to evaluate different data processing techniques to maximize the number of classified liquids. Classic statistics, Discriminant Analysis (DA), and Artificial Neural Networks (ANN) are compared among data processing techniques.

The rest of the paper is structured as follows: Section 2 outlines the related work. The proposal, including the used sensors, node, and architecture of the LIDS system, is described in Section 3. Section 4 details the test bench. The results of the gathered data are analyzed in Section 5. Finally, Section 6 summarises the conclusions and future work.

II. RELATED WORK

In 2016, a study was carried out comparing different ways of handling smart bottles. In this study, an intelligent water bottle was made with the aim of being able to monitor

the consumption of water in a day produced by a person. In order to achieve this, they used sensors of inertia and physiological parameters and a photoplethysmographic sensor. The user could manage this bottle in two different ways: using a smartphone or uploading the data to the cloud. The use of a smartphone is more energy efficient, but you need to be constant during the day, uploading the data periodically. However, the cloud-based system was less energy efficient, but it helps structure many users' data and provides more reliable information [9].

In 2021, research began on the technologies that should be used to make them more efficient and practical for users.

In the first, the LIDS system was studied. In this article, they have focused on creating a system capable of characterizing the type of liquid consumed and its volume. The LIDS provides real-time monitoring of the type and volume of fluid intake. For this, the system designs a detection module comprising ultrasonic, RGB colour, temperature and accelerometer sensors, as well as a computational framework for classifying the type of fluid intake [2].

The second one exposes the available technologies, reviews the existing systems for monitoring the amount of water consumed by users and motivates them to do so by sending notices periodically during the day. It was concluded that the best results are obtained by combining the technologies that are available today: wearable devices, surfaces with integrated sensors, solutions based on vision and the environment, and smart containers [8].

Later, in 2022, the same authors did a study comparing the performance and functionality of four smart bottles on the market. The bottles that were compared were H2Opal, HidrateSpark Steel, HidrateSpark 3 and Thermos Smart Lid. To know the effectiveness of each model, 100 intakes for each bottle were recorded and analyzed, comparing the amount of water consumed using a high-resolution weight scale and the data obtained from the bottle. It was concluded that the best options were the first three, the Smart Lid thermos being the least effective [1]. In this same year, a smart bottle was used for sanitary purposes. The Hydrate Spark bottle was used to count the amount of water ingested by a user prone to stone formation. This bottle had sensors that counted the amount of water in real-time and sent that information to the smartphone, which tells you when you have drunk enough water according to your goal [10].

Among the surveyed proposals, no one of them has focused on developing a system that measures the type of liquid included in the smart bottle. As far as we are concerned, no proposal aimed at classifying the liquids inside the bottles used for LIDS in order to differentiate among intakes of liquids. Thus, this proposal aims to classify the included liquid based on the combination of different patterns of light abortion and include this data in machine-learning solutions, such as DA and ANN.

III. PROPOSAL

This section describes the optical sensor used to detect the different samples in the smart bottle. In addition, the node used and the system's complete architecture are shown.

A. Optical Sensor

This prototype consists of a tube 15 cm high. It is an opaque black PVC tube. In this way, the interference of outside light with the sample is avoided. On the one hand, an RGB LED has been used to differentiate the different test substances, allowing the Arduino programming to establish different wavelengths. In this case, 8 colours have been used: red, green, blue, yellow, purple, cyan and white. The following table, Table 1, shows the colours, intensity and RGB values.

TABLE I. RGB LED INTENSITIES FOR USED COLOURS

Light	Red	Green	Blue
Red	255	255	0
Green	255	0	255
Blue	255	0	0
Yellow	255	255	255
Purple	255	255	0
Cyan	255	0	255
White	255	255	255

On the other hand, the optical sensor contains an infrared LED. The RGB and infrared LED are arranged on the same side. On the opposite side are the visible and IR photoreceptors that make it possible to detect the light that passes through the sample. With this system, it is feasible to detect the absorbance of the tested samples and compare the data between them, which is used to differentiate among tested liquids. The system is based on an enhanced RGB prototype version presented in [11] and [12].

B. Node

An ESP32 microprocessor has been selected to gather and analyze the data. It has been chosen due to its small size, which allows it to be included in the smart bottles in a comfortable way for the user. It has good connectivity and is widely used for Internet of Things (IoT) devices. The different inputs it presents allow the RGB LED to be connected digitally and the Near Infra Red (NIR) LED analogically. The node is presented in Figure 1.



Figure 1. Picture of used node

C. Architecture

The architecture that is proposed for the LIDS-based smart bottle includes different networks. Firstly, the data

obtained from the optical sensor are processed in the node itself by means of the ESP32 microcontroller as part of edge computing. After processing the data, the smart bottle will send the information obtained to a server with a database by connecting with the smartphone using the Bluetooth connection to reach the smartphone and then using the 5G network to reach the database in the server. We have selected using the smartphone and Bluetooth technology in order to ensure connectivity in indoor and outdoor environments regardless of local WiFi connectivity.

Moreover, through a smartphone, the user or the person's caregiver can access the data as a consult to the database. In this way, you can keep track of the liquids ingested by the person as well as the volumes.

When the person is at home, the smart bottle can be connected to a smartwatch or tablet by a WiFi connection. This makes it possible to establish an alarm system that indicates that the user must ingest liquids according to the time that has passed since the last ingestion.

IV. TEST BENCH

This section shows how the samples have been prepared for processing, in addition to how the data has been collected. Finally, the tools used for data processing and subsequent analysis are described.

A. Sample preparation

To develop the intelligent bottle, it used samples for different substances. The processing samples have been: empty, water, tea, Mango tea, light milk, milk, tea with milk and more tea with milk. To prepare the different teas, we introduced the tea to the water. After this, we mix by shaking. The samples are introduced into the intelligent bottle, which is the different colours of LEDs and IR. In the following table (Table 2), we show the characteristics of the samples and the volume included.

TABLE II. RGB LED INTENSITIES FOR USED COLOURS

ID	Sample	Volume (mL)
1	Empty	0
2	Water	50
3	Tea	50
4	Mango tea	50
5	Light milk	50
6	Milk	50
7	Tea with milk	50
8	More tea with milk	50

B. Measuring process

This subsection describes the measuring process. For data collection, the samples are introduced into the smart bottle. Using the SP32, the different colours of RGB LED

are programs. We avoid collecting samples' values when the LEDs and IR illuminate samples. The data are saved like CSV, and after that, are processed.

Each sample is measured 6. In order to generate two datasets, training and validation, the data gathered are split into two groups. The first four replicas were used for the training dataset, while the last two replicas were used for verification.

C. Data processing

Three different approaches to classifying data as a traditional multi-class problem are compared in this paper. First, the ANalises Of Variance (ANOVA) is used as the most simple approach. A series of thresholds are generated based on the results of the multiple groups of the ANOVA. Then, DA and ANN are applied to compare the result of machine learning with simpler methods. The number of wrong classifications in both training and validation datasets is used as a performance indicator.

V. RESULTS

In this section, we are going to present and discuss the obtained results in the calibration and verification of the proposed sensor system for the smart bottle.

A. Descriptive analysis of calibration results

First of all, the average values of the calibration results for each sample and the standard deviation are presented in Figure 2. In general terms, the analogRead() value of the IR light is the largest in terms of deviations and gathered values. The tea is the sample that registered the largest average value and standard deviation, 1460 ± 72 .

In all cases, the higher the value, the greater the amount of light that reaches the photoreceptor. Since milk is a colloidal dilution, the amount of light that it absorbs is greater than that of water. In both water and milk infusions, the addition of the infusions decreases the amount of light that reaches the photodetectors in barely all the lights. In the case of water, this effect is more evident with green, blue, and cyan lights, which might be related to the infusions' pigments.

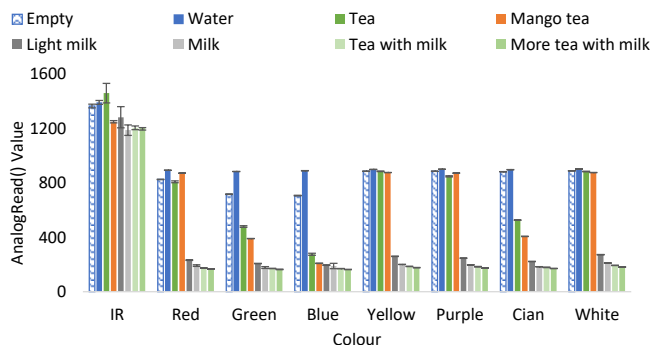


Figure 2. Average and standard deviation of calibration tests.

In order to evaluate which one of the included lights offers better results to assess the classification of drinks included in the bottle, a multivariate analysis is performed to

obtain the correlation matrix. The Sample IDs have been configured according to the values obtained from the analyses of the average data. The results of the correlation matrix with the Pearson product-moment correlation coefficient, see Figure 3, indicated that the lights with the highest correlation with the Sample ID are white, purple, yellow, and green. Thus, these lights will be used for further tests.

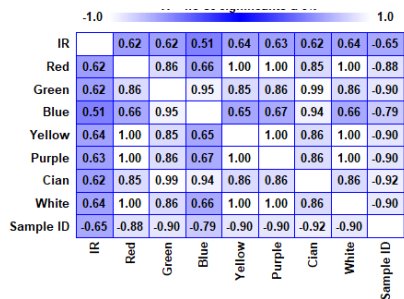


Figure 3. Correlation matrix.

B. Sample classification using thresholds based on ANOVA results

The first classification method tested is the use of a series of thresholds based on the results of the ANOVA and the multiple ranges tests. For this purpose, the results of the ANOVAs for the four selected lights are compared in Table 3. Even with all the lights, the obtained p-value is below 0.001. The multiple ranges tests offered different results for the analyzed lights. The results indicate that the white and purple lights are the ones that correctly divided the samples into individual groups. Using the yellow light and the results of the ANOVA, it is impossible to distinguish Samples IDs 1 and 3, which can provoke that the system is incapable of distinguishing an empty bottle from a bottle with tea. Meanwhile, with the results when the green light is used, the system can confuse Samples IDs 7 and 6 and 7 and 8. It will lead to the incapacity of differentiation between milk with tea and milk or milk with a stronger tea.

Thus, the white and purple lights are selected to be studied in depth in the following subsection. Even though all the lights can be used, the aim of reducing the required lights is to reduce the energy consumption of the bottle in both data gathering and data processing. Table 4 summarizes the

TABLE III. SUMMARY OF ANOVA RESULTS

Light	Sample ID								p-value
	1	2	3	4	5	6	7	8	
White	890.0 ^b	902.0 ^a	883.75 ^c	876.25 ^d	273.0 ^e	211.75 ^f	193.75 ^g	181.5 ^h	<0.0001
Purple	886.25 ^b	900.25 ^a	848.24 ^c	872 ^d	247.25 ^e	197.5 ^f	184.0 ^g	174.5 ^h	<0.0001
Yellow	886.25 ^b	900 ^a	885.25 ^b	877.25 ^c	260.25 ^d	200.25 ^e	186 ^f	177 ^g	<0.0001
Green	717.5 ^b	884.25 ^a	479.5 ^c	390.75 ^d	207.75 ^e	178.75 ^f	171.5 ^{fg}	164.5 ^g	<0.0001

a. Different letters indicate different groups.

threshold values for the classification with the ANOVA results.

To evaluate the degree of efficiency of classification using the ANOVA thresholds, two alternative and well-known methods are evaluated. The classification results with ANN and DA can be seen in Figures 4 and 5. In this case, two different classification approaches have been followed. On the one hand, all gathered data, including all light sources, is used. As in the previous case, the four calibration repetitions are used for the training dataset. On the other hand, a second classification is conducted using only the data of purple and white light according to the results of the ANOVA and multiple ranges tests.

In Figure 4, we can see the confusion matrixes of the training dataset for DA when all lights are used a) and with selected lights b). It is possible to see that there are no differences between both classifications.

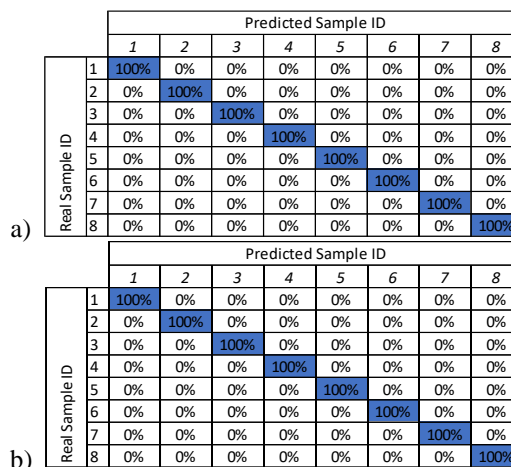


Figure 4. Confusion matrixes with DA when all data is used a) and when selected data is used b).

On the contrary, in Figure 5, the results of the obtained confusion matrixes with ANN when all data is selected a) and only white and purple data are included b) have a great variation. When all lights are used, there is a high classification error in the samples based on milk (including light milk, milk and milk with teas). This error is absent when filtered data, including only the values obtained with the purple and white light.

TABLE IV. TRESHODLS FOR CLASSIFICATION WITH ANOVA RESULTS

Sample ID	White thresholds		Purple thresholds	
	Maximum	Minimum	Maximum	Minimum
1	1024.0	893.4	1024.0	896.1
2	893.3	879.2	896.0	887.0
3	879.1	853.1	886.9	880.1
4	867.2	536.0	880.0	574.7
5	559.6	222.5	574.6	242.5
6	222.4	190.9	242.4	202.9
7	190.8	179.4	202.8	187.7
8	179.3	0.0	187.6	0.0

C. Sample classification using DA and ANN

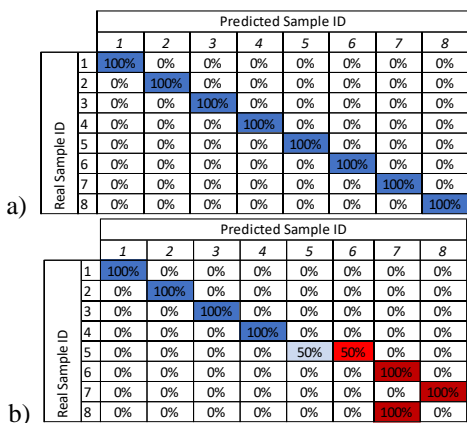


Figure 5. Confusion matrices with DA when all data is used a) and when selected data is used b).

D. Verification

In order to decide which method should be used to determine the content of the bottle, the three presented methods in the classification are compared. Confusion matrices are used to compare the results of the verification.

Figure 6 contains the confusion matrixes using the thresholds obtained with the ANOVA for purple a) and white b) data. It is possible to see that among the individual lights, the white one offered better results. The classification errors are linked to confusing tea with mango tea and milk with milk and tea.

Regarding the results of DA, Figure 7 presents the confusion matrixes of the verification test with all a) and selected data b). The classification is better when only selected data (purple and white combined lights) are used in DA. If all data is used, there are misclassifications between milk and different tea types.

Finally, the verification test results based on ANN with all a) and selected data b) can be seen in Figure 8. Again, the results indicate that the classification is more accurate when only white and purple light data is used. Nevertheless, in this

case, the error is lesser than the errors in the training dataset classification. The error is limited to the sample composed of milk, Sample ID equal to 6, which is confused with milt and tea.

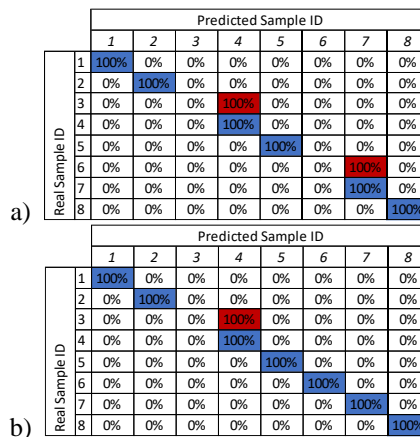


Figure 6. Confusion matrixes of verification test with purple a) and white b) individual data based on ANOVA thresholds.

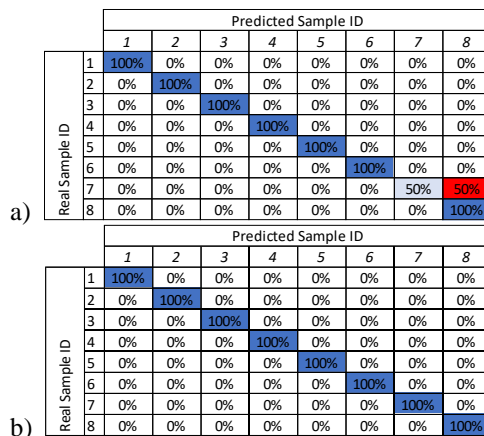


Figure 7. Confusion matrixes of verification test with all a) and selected b) data based DA.

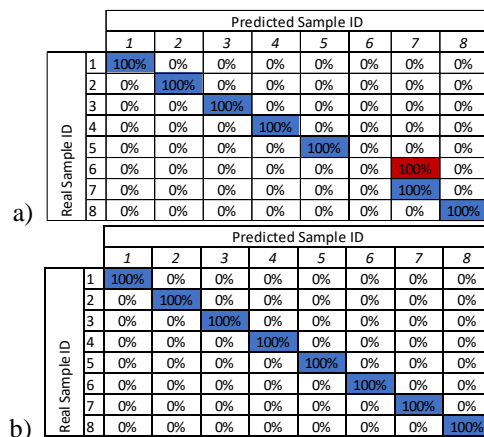


Figure 8. Confusion matrixes of verification test with all a) and selected b) data based ANN.

Thus, we can conclude that the best classification method is the DA. Although the best results are obtained with DA, ANOVA thresholds can be used if there is no need for differentiation between used teas.

VI. CONCLUSIONS

The research seeks to enhance current liquid intake detection systems by introducing an RGB sensor for discriminating liquids, facilitating precise hydration assessment. Different approaches for data classification are compared.

Results demonstrate successful differentiation using a limited number of light wavelengths, primarily white and purple. Discriminant Analysis stands out as the optimal classification method, with 100% of cases correctly classified in the verification phase. Although ANOVA thresholds can be used to achieve similar results with less computational demand when differentiating teas isn't essential.

Future work involves including additional sensors in the smart bottle to detect abnormalities in the liquids. Moreover, the integration of the smart bottle with other devices, creating an IoT solution for ambient assisted living and eHealth for elderly people, is foreseen as part of the ongoing projects.

ACKNOWLEDGMENT

This work is partially funded by the "Programa Estatal de I + D + i Orientada a los Retos de la Sociedad, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020 project PID2020-114467RR-C33/AEI/10.13039/501100011033, by "Proyectos Estratégicos Orientados a la Transición Ecológica y a la Transición Digital" project TED2021-131040B-C31, and by the ThinkInAzul programme and was supported by MCIN with funding from European Union NextGenerationEU (PRTR-C17.I1) and by Generalitat Valenciana (THINKINAZUL/2021/002).

REFERENCES

- [1] R. Cohen, G. Fernie, and A. Roshan Fekr, "Monitoring fluid intake by commercially available smart water bottles," *Scientific Reports*, vol. 12, no. 1, pp. 4402, 2022.
- [2] M. Pedram, et al. "Lids: Mobile system to monitor type and volume of liquid intake," *IEEE Sensors Journal*, vol. 21, no. 18, pp. 20750-20763, 2021.
- [3] N. Pushpalatha et al. "A Review on Intelligent Water Bottle Powered by IoT,". In *Proceedings of the 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 17-18 March 2023, pp. 2335-2338.
- [4] S. Flutura et al. "Drinkwatch: A mobile well-being application based on interactive and cooperative machine learning". In *Proceedings of the 2018 International Conference on Digital Health (DH'18)*, Lyon, France, 23-26 April 2018, pp. 65-74.
- [5] M. Rosenberg, et al. "Wearable alcohol monitors for alcohol use data collection among college students: feasibility and acceptability," *Alcohol*, vol. 111, pp. 75-83, 2023.
- [6] M. S. Borofsky, C. A. Dauw, N. York, C. Terry, and J. E. Lingeman, "Accuracy of daily fluid intake measurements using a "smart" water bottle," *Urolithiasis*, vol. 46, pp. 343-348, 2018.
- [7] L. B. Baker, et al. "Accuracy of a smart bottle in measuring fluid intake by American football players during pre-season training," *Scientific Reports*, vol. 13, no. 1, pp. 11383, 2023.
- [8] R. Cohen, G. Fernie, and A. Roshan Fekr, "Fluid intake monitoring systems for the elderly: A review of the literature," *Nutrients*, vol. 13, no. 6, pp. 2092, 2021
- [9] E. Jovanov, V. R. Nallathimareddygar, and J. E. Pryor, "SmartStuff: A case study of a smart water bottle," In *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'16)*, Orlando, USA, 16-20 August 2016, pp. 6307-6310
- [10] T. E. Stout, et al "A randomized trial Evaluating the Use of a smart water bottle to increase fluid intake in stone formers," *Journal of Renal Nutrition*, vol. 32, no. 4, pp. 389-395, 2022.
- [11] S. Sendra, L. Parra, V. Ortuño, and J. Lloret, "A low cost turbidity sensor development," in *Proceedings of the Seventh International Conference on Sensor Technologies and Applications (SENSORCOMM 2013)*, Barcelona, Spain, 25-31 August 2013, pp. 25-31
- [12] L. Parra, L., S. Viciano-Tudela, D. Carrasco, S. Sendra, and J. Lloret, "Low-cost microcontroller-based multiparametric probe for coastal area monitoring," *Sensors*, 23(4), p. 1871, 2023.