Classification of coffee and wine with a microwave resonator and deep learning machine technique

Hyobong Hong[a,]* Jae-Chan Jeong[a], Hans-Joachim Krause[b],

[a] SW Contents Research Lab., Electronics and Telecommunications Research Institute (ETRI), 218 Gajeong-Ro, Yuseong-Gu, Daejeon 34129, Republic of Korea

[b] Institute of Biological Information Processing, Bioelectronics (IBI-3), Forschungszentrum Jülich, 52425 Jülich, Germany

[Abstract]

In this study, coffee and wine were measured using an microwave resonator, and a deep learning system was trained using the acquired data, and then tested to see if the deep leaning system could distinguish these samples. We tested 6 kinds of wine, 6 kinds of cold brew coffee and 6 kinds of bottled coffee. The microwave resonance spectra of all samples were graphically displayed. The graphical images were processed by an artificial intelligence (AI) technique. By applying deep learning machine technique instead of the peak assignment for complex compounds in general, it was possible to facilitate the classification of coffee or wine with high accuracy. We believe that the new analytical method presented in this paper will be very useful for tracking changes in very complex chemical components such as food and natural products.

[Introduction]

Many people habitually drink coffee after waking up in the morning and during their daily work. Also, in the evening, many people enjoy wine with meals or to add to the joy of the party. Quite often, there is a particular type of coffee that an individual enjoys most, and a wine that is preferred. In fact, the main ingredients that make up coffee or wine do not differ much between products. However, the thousands of compounds contained in small amounts compared to the few main ingredients make a fine but distinct taste difference. This distinct taste difference is just a preference for the average consumer, but maintaining this taste, which is difficult for the producers who make these products, means long experience and a major know-how in the manufacturing process.

It is not easy to maintain a certain quality because about thousand kinds of compounds are known that determine the aroma of coffee, and wine production is the result of very complex chemical reactions during fermentation and aging. In order to maintain quality, the most important step is to measure the difference between a newly produced product and the long-standing reference product. Therefore, most manufacturers rely on sensory testing of artisans with long skill and experience. However, in the case of sensory tests, it is difficult to obtain such a master, and it is natural that the ability of the tests also differs from person to person. For example, not all sommeliers produce the same results in blind wine tests. The problem is that there is currently an objective and quantitative way to record these changes in quality in the "scoring scorecard" [1]. Although chemical analysis methods are widely used, in most cases, the presence or absence of a specific substance is measured, so it is not easy to detect the change in products composed of complex compounds such as coffee or wine[2-3]. In fact, analyzing all the individual peaks in such an analysis does not make much sense. This is because many peaks are generated in one chemical substance, whether it is optical spectroscopy or mass spectrometry, and these peaks overlap and are buried in an analysis in which numerous chemical substances are mixed. Therefore, studies using the "profile" method have been conducted to solve this problem. Primarily, it involves placing the sample in analytical equipment with minimal or no purification[4].

There are several ways to produce data with this complex pattern in coffee or wine analysis (e.g. GC or LC mass spectrometry, spectroscopic analysis using IR or UV, NMR, etc.). However, in order to use GC or LC mass spectrometry, molecular weight limitations are obvious in most cases. In order to use MALDI TOF-mass spectrometry, there are limitations in ionizing all of the various substances that make up food into several matrices. In the case of IR, there is a problem of moisture (unfortunately, both coffee and wine mostly contain moisture) and the requirement that IR must pass through a sample. UV-VIS spectrometry also has a problem that the amount of sample is limited and that the UV must pass through the sample. NMR is a very good data generation equipment, but it also has the disadvantage that the equipment is expensive and more difficult to operate. Therefore, in this study, we decided to acquire the necessary data using a microwave resonator. Microwave measurement techniques using a resonator have the advantage that it is possible to secure very stable data, build an analysis system at a relatively inexpensive price (If the sweeping speed is not an issue, the system can be built for less than 600 Euros.), and control the resolution such as the number of data points per frequency range more accurately than for optical equipment[5]. The system used in this experiment was able to secure a considerable amount of data for each sample. The problem was to distinguish the actual product using this data. Of course, many researchers will be able to find out the characteristic points by distinguishing all the many peaks one by one. Instead, we decided to use deep learning techniques in this study. As a result, it was possible to save a lot of manpower and time by using the developed system. Based on these results, the system we developed showed the possibility of being an effective tool in profiling analysis of complex data.

[Materials and Methods]

In this study, 6 types of cold-brew coffee, 6 types of bottled coffee and 6 types of wine were used. The cold brew coffee used in this study is coffee that is supplied to consumers by using a cold chain without any other pre-treatment after extracting the ground coffee beans using cold water. Consumers usually dilute this coffee three times at home. Bottled Coffee is coffee that comes in a plastic bottle so that consumers can drink it without brewing. The information on the samples used is summarized in Table 1.

Table 1 . List of the coffees and wines tested in this study

	Brand name of Coffee	Manufacturer	Abbreviation
Cold brew Coffee	Kenya AA	1kg Coffee (Seoul, Korea)	AA
	Cost Rica Natural	1kg Coffee	CN
	Indonesia Mandheling	1kg Coffee	G1
	Ethiopia Yirgacheffe	1kg Coffee	G2
	Decaffeina Blending	1kg Coffee	Joy
	Blending Seolem	1kg Coffee	SL
Liquid Coffee	Contra base Black	Lotte, Co. LTD (Seoul, Korea)	Base
	Cantata Black	Lotte, Co. LTD (Seoul, Korea)	Cantata
	Colombia Master Black	Heinz, Co. LTD (U.S)	CMB
	Craft Black	Coca-Cola (Japan)	GCB
	Gotica Black	Coca-Cola (Japan)	JGT
	T.O.P Black	Dong-Suh Co. LTD (Seoul, Korea)	TOP

	Manufacturer	Grape varieties	Origin	Vintage	Abbreviation
Wine	Carlo Rossi	California Red	U.S	N/A	CRC
	Carlo Rossi	Concord	U.S	N/A	CRCG
	Gato Negro	C. Sauvignon	Chile	N/A	GNC

Legend of Chile	C. Sauvignon	Chile	2019	LCS
Legend of Chile	Merlot	Chile	2019	LMER
Santa Helena	Carmenere	Chile	2018	SHVC

Of all products, 50ml were used without additional pretreatment. As sample holder inside the microwave resonator, a commercially available plastic container (450.0 ml) was used. As a negative reference, the container was filled with 50.0 ml of distilled water. It was measured 10 times while changing the water or changing the position of the sample. Of all the samples, three bottles of the same type were purchased. Three samples per bottle (150.0 ml in total) were taken. After collecting, each sample was measured 10 to 12 times, but once again every 2-3 times, the sample was changed or shaken slightly before measuring again. The reason for doing this is to include the minute changes in the signal during the actual measurement. The microwave resonator used in this experiment was originally designed for moisture determination. Details are mentioned in the papers published [6-7]. For the basic design and verification of the antenna, an on-line design tool was used [8].

Briefly, the microwave resonator is a cylindrical double-cone-shaped aluminum structure with sleeves determining the inner diameter, see the photograph at the left of Figure 1. The parameters and calculated characteristics used in the fabrication of this microwave resonator can be found in Table 2.

Table 2. Parameter and values of Microwave resonator and antenna

	Parameter	Value
Resonator design parameter	Resonator inner diameter	110 mm
	Resonator outer diameter	275 mm
	Resonator height	148 mm
	Resonance of TM_{010} mode[7]	1150 MHz
Selected antenna design parameter	Antenna diameter	15 mm
	Loop antenna wire length	49 mm
	Conductor diameter	0.6 mm
	Transmitter power	1.0 mW

In this study, a Vector Network Analyzer (ZNLE 6, Rohde & Schwarz, München, Germany) was used to excite the resonator with the sample at one loop antenna and secure the signal at the receiving loop antenna (S21 mode). The antennas for transmitting and receiving signals are mounted at 90° angle to each other, as described in Ref. 7. For calibration of VNA, ZV-Z135 of the same company was used. The parameter setting of the equipment is as follows: Span: 1.0MHz ~ 6.0GHz, Reference D / B: 10.0, Sweep: 5,000 points, S21 mode. Thus, the broadband electromagnetic transmission through the resonator-sample system was measured. As described in Ref. 7, several different transversal-magnetic (TM) modes are excited in the resonator. Their center frequencies and quality factors are influenced by the dielectric permittivity $\varepsilon(f)$ and by the electrical conductivity $\sigma(f)$ of the sample at the particular mode's frequency f. If the conductivity of the liquid sample is high, the electric field at its border is strongly suppressed. Depending on the type of mode, this may lead to a deformation or suppression of the mode. Thus, the frequency-dependent complex dielectric permittivity of the sample influences its measured S21 transmission spectrum. Care must be taken that the position of the sample container is kept unchanged for all measurements. As mentioned above, the acquired data were converted into graphs using our homemade software. The developed software has been uploaded to Github (https://github.com/jaechanjeong/CVSParser.git). This software can not only process data generated by Rohde & Schwarz network analyzers but also from cheaper equipment such as Signal Hound's spectrum analyzer with tracking generator (Battle Ground, WA, US). All data used in this paper were acquired using Rohde & Schwarz products. The function of our software is to automatically read the CVS format data obtained in the experiment, calculate the magnitude value using the data divided into real and imaginary values, and subtract the reference data from water. The secured image was used to train an on-line deep learning machine, a teachable machine (https://teachablemachine.withgoogle.com/). Then, model export was performed for Tensorflow. Tensorflow is a symbolic math library that is a data flow programming used in AI applications [9]. Classification was performed on a PC with Python 3.0 installed or directly on the teachable-machine home page. The following diagram briefly shows how the data was processed and classified in this experiment (Fig. 1).

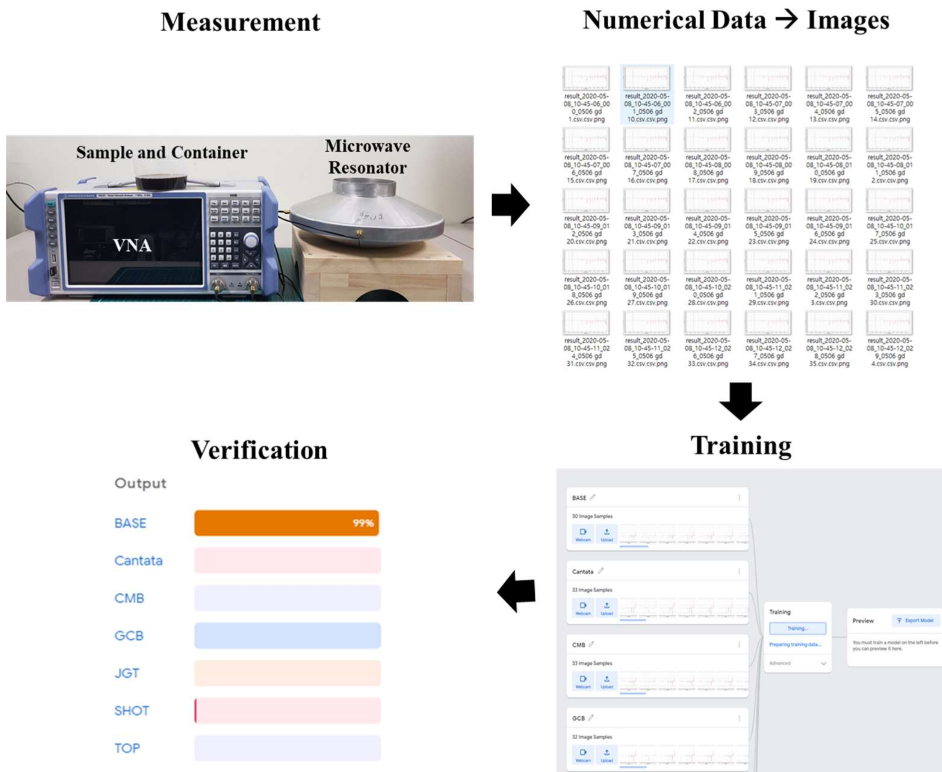

Fig 1. The workflow employed in this research.

[Results and Discussion]

In this study, after the secured data was graphically displayed, the artificial intelligence system was trained using the image of this graph, and verification of the system was performed. The reason for adding the conversion step from numerical data to graphical image is that if the generated data is used directly, the process of conversion of the data format for each device is cumbersome. Also, in the case of a deep learning processing system that can be used in general (at least for people who are not AI experts), the image has the advantage that it can be used immediately without any data processing. As a result of the experiment, there was no problem in distinguishing the product type with the system built in this study. For comparison, we also tried to analyze the experimental data directly by comparing the magnitude values at the recorded 5,000 frequencies per sample one by one. However, it turned out that it is almost impossible to analyze the data this way.

Table 3 shows the results for the materials used in this study. As shown in the experimental results, all six types of coffee were accurately classified. However, the measurement probability was different for each coffee. In the case of one brand name (Cost Rica Natural), it appeared that there was no similarity with other coffees used in the experiment. No other coffees have any similarities to CN. On the other hand, in the case of G1 (Indonesia, Mandheling), it was found that there were some similarities with other coffees under these experimental conditions.

Table 3. Results of the classification of the cold brew coffee

		AA	**CN**	**G1**	**G2**	**JOY**	**SL**
AA	Average	82.60	0.00	2.20	6.60	1.60	7.00
	STDEV	9.24	0.00	1.79	5.32	3.05	4.00
CN	Average	0.00	100.00	0.00	0.00	0.00	0.00
	STDEV	0.00	100.00	0.00	0.00	0.00	0.00
G1	Average	12.60	0.00	51.80	8.20	20.60	6.80
	STDEV	9.10	0.00	7.29	3.70	8.44	2.86
G2	Average	6.40	0.00	6.00	70.60	7.80	10.00
	STDEV	3.85	0.00	4.18	14.06	7.82	2.83
JOY	Average	1.20	0.00	3.80	2.20	84.20	8.60
	STDEV	0.84	0.00	1.30	0.84	5.89	4.51
SL	Average	12.00	0.00	1.40	7.80	8.20	70.60
	STDEV	8.49	0.00	0.55	8.87	7.36	12.16

Table 4 shows the results of different bottled coffee sold on the market. Experimental results show that there is a clear difference between products. It can be concluded that there are more differences as compared to the results in Table 3 based on products produced by the same company. However, the aim of this paper is not to discuss the analysis of a specific product, but to demonstrate the reliability of distinction between different products.

Table 4. Results of the classification of the bottled coffee

		BASE	CANTATA	CMB	GCB	JGT	TOP
BASE	Average	99.80	0.40	0.00	0.00	0.00	0.00
	STDEV	0.40	0.50	0.00	0.00	0.00	0.00
CANTATA	Average	0.00	96.80	3.20	0.00	0.00	0.00
	STDEV	0.00	1.64	1.64	0.00	0.00	0.00
CMB	Average	0.00	1.20	98.80	0.00	0.00	0.00
	STDEV	0.00	0.40	0.40	0.00	0.00	0.00
GCB	Average	0.00	0.00	0.00	99.60	0.00	0.00
	STDEV	0.00	0.00	0.00	0.54	0.00	0.00

JGT	Average	0.00	0.00	0.00	0.00	100.00	0.00
	STDEV	0.00	0.00	0.00	0.00	0.00	0.00
TOP	Average	0.00	0.00	0.00	0.00	0.00	99.00
	STDEV	0.00	0.00	0.00	0.00	0.00	100.00

Table 5 shows the results obtained for wine. Like coffee, it was possible to accurately classify the samples by product. Of course, as with coffee, for verification of the system, a different data set than that for system training was used.

Table 5. Results of the classification of the selected wine

		CRC	CRCG	GNCS	LMER	LCS	SHVC
CRC	Average	100.00	0.00	0.00	0.00	0.00	0.00
	STDEV	0.00	0.00	0.00	0.00	0.00	0.00
CRCG	Average	0.00	100.00	0.00	0.00	0.00	0.00
	STDEV	0.00	0.00	0.00	0.00	0.00	0.00
GNCS	Average	0.00	1.20	99.40	0.60	0.00	0.00
	STDEV	0.00	0.40	0.40	1.34	0.00	0.00
LMER	Average	0.00	0.00	0.80	99.20	0.00	0.00
	STDEV	0.00	0.00	0.83	0.83	0.00	0.00
LCS	Average	0.00	0.00	0.00	0.00	100.00	0.00
	STDEV	0.00	0.00	0.00	0.00	0.00	0.00
SHVC	Average	0.00	0.00	6.80	0.00	0.00	93.00
	STDEV	0.00	0.00	8.34	0.00	0.00	8.51

In this study, different types of coffee and wine were examined by using a microwave-based measurement technology. The technology has been used for a long time for determining water content, but has been rarely used for the analysis of complex organic matter, and in conjunction with relatively novel deep learning technology [10]. Deep learning technique is used in many fields, from playing Go with humans to autonomous vehicles in these days [11].

The biggest advantage of deep learning is that training and classification can be done at a much faster rate than humans can do. In this regard, the deep learning technique in this study was a very good assistant. It was very difficult for the researchers to review dozens or hundreds of samples with 5000 data points per sample, remember the pattern, and then classify it when a validation sample comes in. However, there were also obvious disadvantages of deep learning. It has been very difficult to analyze which components or peaks of a specific product determine these characteristics in a manner developed to date. The solution to this problem will be secured through research using the samples with well-known ingredients.

As a result of this study, we want to highlight two things. If one wants to classify substances composed of very complex compounds, any type of analytical device that can generate various kinds of signals according to the properties of the sample (for instance a microwave measurement of the resonance properties of the samples influenced by their frequency-dependent dielectric permittivity, as done in this study) in addition to the existing analytical chemistry methods can be combined with deep learning techniques. However, with the current technology, it is only possible to measure the degree of change of a certain parameter. In further studies, one could try to analyze exactly which component or ingredient leads to this change.

[Acknowledgement]

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No. 2015-0000-20

[References]

1. Barbosa, M. d. S. G.; Scholz, M. B. d. S.; Kitzberger, C. S. G.; Benassi, M. d. T., Correlation between the composition of green Arabica coffee beans and the sensory quality of coffee brews. *Food Chemistry* **2019,** *292*, 275-280.

2. Bressanello, D.; Liberto, E.; Cordero, C.; Rubiolo, P.; Pellegrino, G.; Ruosi, M. R.; Bicchi, C., Coffee aroma: Chemometric comparison of the chemical information provided by three different samplings combined with GC–MS to describe the sensory properties in cup. *Food Chemistry* **2017,** *214*, 218-226.

3. Farah, A.; Donangelo, C. M., Phenolic compounds in coffee. *Brazilian Journal of Plant Physiology* **2006,** *18*, 23-36.

4. Sobolev, A.; Mannina, L.; Aru, V.; Bellomaria, A.; Bertocchi, F.; Botta, B.; Cagliani, L.; Caligiani, A.; Capozzi, F.; Çela, D.; Marincola, F.; Ciampa, A.; Del Coco, L.; Consonni, R.; Corsaro, C.; Delfini, M.; Di Tullio, V.; Fanizzi, F.; Gallo, V.; Capitani, D., NMR applications in food analysis: Part A. 2017.

5. Jones, G. E.; Cook, R. L.; Lide, D. R., Chemical Analysis by Microwave Spectroscopy. *C R C Critical Reviews in Analytical Chemistry* **1973,** *3* (4), 455-506.

6. MENZEL, M. I.; TITTMANN, S.; BÜHLER, J.; PREIS, S.; WOLTERS, N.; JAHNKE, S.; WALTER, A.; CHLUBEK, A.; LEON, A.; HERMES, N.; OFFENHÄUSER, A.; GILMER, F.;

BLÜMLER, P.; SCHURR, U.; KRAUSE, H.-J., Non-invasive determination of plant biomass with microwave resonators. *Plant, Cell & Environment* **2009,** *32* (4), 368-379.

7. Sydoruk, V.; Fiorani, F.; Jahnke, S.; Krause, H.-J., Design and Characterization of Microwave Cavity Resonators for Noninvasive Monitoring of Plant Water Distribution. *IEEE Transactions on Microwave Theory and Techniques* **2016,** *64*.

8. 66pacific.com. http://www.66pacific.com/calculators/smalltransmitting-loop-antenna-calculator.aspx.

9. Mart´ın Abadi, A. A., Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro,; Greg S. Corrado, A. D., Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow,; Andrew Harp, G. I., Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser,; Manjunath Kudlur, J. L., Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, ´; Chris Olah, M. S., Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar,; Paul Tucker, V. V., Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, ´; Pete Warden, M. W., Martin Wicke, Yuan Yu, and Xiaoqiang Zheng; Research∗, G., TensorFlow:

Large-Scale Machine Learning on Heterogeneous Distributed Systems. 2015.

10. LeCun, Y.; Bengio, Y.; Hinton, G., Deep learning. *Nature* **2015,** *521* (7553), 436-444.

11. Silver, D.; Huang, A.; Maddison, C. J.; Guez, A.; Sifre, L.; van den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; Dieleman, S.; Grewe, D.; Nham, J.; Kalchbrenner, N.; Sutskever, I.; Lillicrap, T.; Leach, M.; Kavukcuoglu, K.; Graepel, T.; Hassabis, D., Mastering the game of Go with deep neural networks and tree search. *Nature* **2016,** *529* (7587), 484-489.