Preprocessing Taste Data for Deep Neural Networks

Hyunjong Lee
Department of Electrical Engineering and Computer Science
Daegu Gyeongbuk Institute of Science and Technology (DGIST)
Daegu , Republic of Korea
hjlee@dgist.ac.kr

Han Hee Jung
Department of Robotics and Mechatronics Engineering
DGIST
Daegu , Republic of Korea
hanheejung18@dgist.ac.kr

Junwoo Yea
Department of Robotics and Mechatronics Engineering
DGIST
Daegu , Republic of Korea
yehjunwoo@dgist.ac.kr

Kyung-In Jang
Department of Robotics and Mechatronics Engineering
DGIST
Daegu, Republic of Korea
kijang@dgist.ac.kr

Jeongho Kwak
Department of Electrical Engineering and Computer Science
DGIST
Daegu, Republic of Korea
jeongho.kwak@dgist.ac.kr

Jihwan P. Choi
Department of Aerospace Engineering
Korea Advance Institute of Science and Technology (KAIST)
Dajeon, Republic of Korea
jhch@kaist.ac.kr

***Abstract*— Analyzing wine using taste data is a promising field due to the explosive expansion of online commerce. However, because of the wide variety of wine types with different flavors and aromas, it is difficult for consumers to choose the wine that suits their taste, and also difficult for sellers to recommend appropriate wines to consumers. Therefore, it is necessary to numerically analyze and classify wine, and a deep learning algorithm which mimics the human brain is appropriate for analyzing the wine data [1]. In this paper, we introduce several studies of wine classification using deep learning architectures and propose preprocessing methods for applying the taste data of wine to deep learning networks.**

Keywords—wine, taste sensor, taste classification, deep learning, preprocessing, data augmentation

I. Introduction

Wine is one of the most popular alcoholic beverages in the world and has many health benefits [2]. The popularity of wine lies in its unique and diverse taste. Depending on the variety of grapes, wines with various tastes and aromas can be made, and even with the same variety of grapes, the taste of wine varies depending on the growing season or the place of origin of the grapes [3]. However, it is not easy for general consumers to taste various types of wine, and it is not easy to figure out the characteristics of wine by reading the descriptions of wine. These days, with the explosive growth of the online commerce market [4], it is becoming more difficult to taste and buy wine. For this reason, the analysis of wine data is becoming important for the application of wine classification and personalized recommendation. In this paper, we introduce studies using deep learning as a method of classifying wine and extracting features, propose pre-processing methods for taste sensor data, and introduce several machine learning techniques accordingly.

II. Relaed Work

There are commonly two types of studies that try to classify wine or extract features from the wines. The first one focuses on the chemical properties of wine. These studies classify wine using values of chemical components of wine such as potential of hydrogen (pH), fixed acidity, total sugar or alcohol ratio. Using such data as input has the advantage of being relatively easy to obtain, but it is quite different from the actual taste that human feel.

The other is the case of using the taste of wine measured by electronic sensors as an input. In that case, because they measure sweet, salty, sour, and bitter tastes, just like the human sense of taste, it has the advantage that various applications such as personalized recommendation in the online market are possible.

A. Wine Classification Using Chemical and Physical Features as Data

In [5] the quality of wine is predicted by using alcohol, pH, sulphates and free sulfur dioxide as feature data. The authors of the paper use the methods of Ridge Regression, Support Vector Machine (SVM), Gradient Boosting Regressor (GBR), Artificial Neural Network (ANN) with four-layers, confirming that the ANN performs better than other mathematical models in most of the datasets. There is also a study that predicts the taste of rice wine using deep neural networks with inputs of alcohol, total acid, total sugar, and sugar-removing solid [6]. Some studies used the physical characteristics of wine as input data rather than the chemical characteristics. The authors of [7] used the Raman spectroscopy, which uses properties of scattering light, to classify wine. The Raman spectra of red wine were used as input data for four different types of deep learning networks, ANN, multifeature fusion convolutional neural network (MCNN), GoogLeNet, and residual neural network (ResNet). In [8], hyperspectral imaging was used as input for deep learning networks to predict the pH and sugars of wine grapes.

B. Classification Using Taste Data

All of the studies introduced earlier are far from the characteristics that human actually sense. There are some studies that try to measure the taste that a person senses [9-11]. Since the taste of all foods, not just wine, gives complex stimuli to the electronic sensors, a deep learning network can be used to process these data. However, methods of analyzing these data for deep learning networks have not been much studied to the best of our knowledge. The authors of [12] used ANN to classify wines stored in various environments. In this paper, we describe the characteristics of these taste data that are measured by taste sensors and propose preprocessing techniques to be used as inputs for deep learning networks.

III. CHARACTERISTICS OF TASTE DATA AND PREPROCESSING

A. Characteristics of Taste Data

In order to mimic the human gustatory system, food or beverage must be quantified and measured. It is known that more than four types of tastes are perceived by humans: sweet, salty, sour, bitter and so on. Therefore, to generate taste data, multi-channel sensors capable of measuring multi tastes are required.

However, taste data measured by multi-channel taste sensors have several limitations different from other human sensory data such as image data or auditory data. First, there is a limit to mass production. It takes a long time to measure individual taste data compared to other types of data because it has to wait for the food to spread through the solution, after which it takes time for the sensor to initialize again. Thus, it takes too much time for mass production. Second, the performance of the sensors also declines according to time, and the food or beverage data also change over time [9]. In particular, in case of wine, oxidation of wine proceeds from the moment the wine bottle is opened. To overcome this limitation of taste data, preprocessing for data and data augmentation methods for deep learning are proposed.

B. Data Preprocessing for Neural Networks

In order to use the taste sensor data of wine as an input value to the deep learning network, data preprocessing was performed under three conditions. The first is to remove outlier data to improve the weakness of the taste sensor and wine. As mentioned above, the sensor's performance declines over time and the probability of producing error data increases. Wine can be also oxidized or contaminated when taste is measured with taste sensors, so that error data can be generated. For these reasons, it is necessary to remove data when the value differs greatly from others.

The second is to make the input data of deep learning similar to the human's gustatory system. In order to imitate the taste that a person feels, it is necessary to generate one input data by binding multiple tastes that the human gustatory system senses. However, depending on the lifespan of the sensor, there may be differences in the number of taste data within even one wine. When the difference in the number of data is severe, a Gaussian noise is added to the data to balance the amount of data between the tastes of each type of data.

The third is to secure enough input data for deep learning networks. As mentioned above, current taste sensors are difficult to produce enough data to train deep learning networks. For this reason, it is necessary to use an appropriate data augmentation method for taste data.

In order to satisfy the above three conditions, the first step is to normalize the sensor data and remove sensor data that are outside the threshold. Through this process, outliers in the sensor data are removed. The next step is to collect sensor data by the taste of each type of wine and name the collected data as Taste Bank. Through the combination of tastes in this Taste Bank, data in a form similar to what the human's gustatory system senses is created. This process also has the advantage of increasing the amount of data. Since one taste data is created by selecting one taste from each of the four tastes and combining them, the number of data can be increased. For example, if the electronic taste sensors measure sweet, salty, sour, and bitter tastes, let us denote each number of sweet, salty, sour, and bitter taste data as N_{sweet}, N_{salty}, N_{sour}, and N_{bitter}, respectively. Then, the number of taste input data that can be generated with these

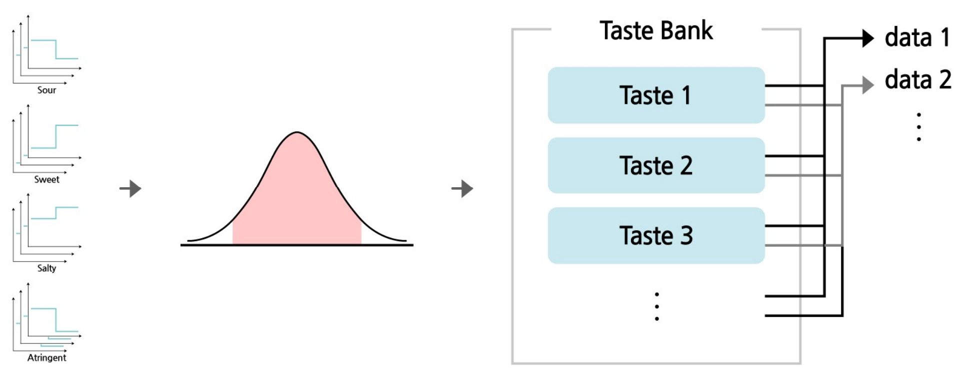

Figure 1. Preprocess for taste data

sensor data is $N_{sweet} \times N_{salty} \times N_{sour} \times N_{bitter}$. The above process is illustrated in Figure 1. Through these steps, the above-mentioned limitations of the taste sensor data can be solved to some extent.

IV. CONCLUSION

In this paper, various studies for wine classification are introduced and the characteristics of taste data are explained. The methods for preprocessing and data augmentation for deep learning networks are also proposed. Through the proposed preprocessing methods, it is possible to transform the sensor data to dataset which is similar to the human's gustatory system. It is also possible to generate more data that are necessary for training machine learning algorithms or deep learning networks by greatly increasing the number of insufficient data.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) funded by Ministry of Science and ICT (NRF- 2019M3C1B8090845, 2020R1C1C1013030, RS-2023-00234581, 23-SENS-01)

REFERENCES

[1] Dong, Shi, Ping Wang, and Khushnood Abbas. "A survey on deep learning and its applications." Computer Science Review 40 (2021): 100379.

[2] Ursini, Fulvio, and Alex Sevanian. "Wine polyphenols and optimal nutrition." Annals of the New York Academy of Sciences 957.1 (2002): 200-209.

[3] Novikova, Liubov Yu, and Lyudmila G. Naumova. "Dependence of fresh grapes and wine taste scores on the origin of varieties and weather conditions of the harvest year in the northern zone of industrial viticulture in Russia." Agronomy 10.10 (2020): 1613.

[4] Akter, Shahriar, and Samuel Fosso Wamba. "Big data analytics in E-commerce: a systematic review and agenda for future research." Electronic Markets 26 (2016): 173-194.

[5] Dahal, K. R., et al. "Prediction of wine quality using machine learning algorithms." Open Journal of Statistics 11.2 (2021): 278-289.

[6] Li, Jian, et al. "Deep learning based flavor evaluation of rice wine." AIP Conference Proceedings. Vol. 2208. No. 1. AIP Publishing, 2020.

[7] Lu, Bingxu, et al. "Identification of Chinese red wine origins based on Raman spectroscopy and deep learning." Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 291 (2023): 122355.

[8] Gomes, Véronique, Ana Mendes-Ferreira, and Pedro Melo-Pinto. "Application of hyperspectral imaging and deep learning for robust prediction of sugar and pH levels in wine grape berries." Sensors 21.10 (2021): 3459.

[9] Wu, Xiao, et al. "Taste sensor: Electronic tongue with lipid membranes." analytical sciences 36.2 (2020): 147-159.

[10] Zhu, Yanchao, et al. "A fluorescent sensor array-based electronic tongue for Chinese tea discrimination." Journal of Materials Chemistry C 9.17 (2021): 5676-5681.

[11] Khan, Amir, et al. "Self-healable and anti-freezing ion conducting hydrogel-based artificial bioelectronic tongue sensing toward astringent and bitter tastes." Biosensors and Bioelectronics 198 (2022): 113811.

[12] Riul Jr, Antonio, et al. "Wine classification by taste sensors made from ultra-thin films and using neural networks." Sensors and Actuators B: Chemical 98.1 (2004): 77-82.