2025 IEEE/SICE International Symposium on System Integration (SII)
January 21-24, 2025, Munich, Germany

Extraction of color information array from RGB-NIR images enhanced by multispectral illumination and image classification by LLGMN

Taiga Eguchi[1], Wen Liang Yeoh[1], Hiroshi Okumura[1] and Osamu Fukuda[1]

Abstract— **In recent years, advancements in image classification technology have led to significant improvements in classification accuracy. However, challenges remain, such as difficulties in image classification when redundant information is present, even when using state-of-the-art deep learning methods, and the need for large amounts of training samples for deep learning models. To address these issues, we proposes a method that enhances color information for image classification by combining multi-illumination and multispectral cameras, and utilizes log-linearized gaussian mixture neural network that can classify images with a small number of training samples. The proposed system utilizes a multi-spectral camera capable of capturing Red (R), Green (G), Blue (B), and Near-Infrared (NIR) images (RGB-NIR), along with corresponding multi-spectral illumination. By enhancing color information and clarifying differences in object features, this approach enables high-accuracy color classification with an unprecedentedly small dataset when input into the log-linearized gaussian mixture neural network. Experimental results demonstrated the effectiveness of the proposed system, achieving 100% color classification accuracy on green tea samples with similar color features. This achievement is expected to contribute to various fields such as manufacturing, healthcare, chemistry, and agriculture, where multispectral imaging is increasingly utilized.**

I. INTRODUCTION

Image classification, which is in high demand across various fields such as manufacturing, healthcare, and agriculture, has experienced rapid technological advancements, largely due to the development of numerous datasets, including ImageNet [1]. Research in this area has shown that enhancing color information plays a crucial role in achieving high-precision image classification using machine learning and deep learning techniques [2]. By enhancing color information before inputting it into machine learning or deep learning models, solutions to various classification challenges have been sought.

Methods for enhancing color information can be broadly categorized into two approaches: using multi-spectral image sensors and enhancing the information from acquired images. In image sensing, color information enhancement can be achieved by using multispectral cameras to capture images across multiple wavelength bands. Specifically, four-sensor line-scan cameras, which can simultaneously capture images in the visible Red (R), Green (G), and Blue (B) bands as well as the near-infrared (NIR) band [3], have been widely utilized in manufacturing, agriculture, and healthcare fields [4]–[6]. By combining the information from the RGB spectral band (400-700 nm) captured by conventional RGB color cameras with the NIR spectral band (700-1100 nm) that is imperceptible to the human eye, rich image features can be obtained, contributing to improved image classification accuracy. Recently, datasets from these application fields have been made public, and image classification algorithms tailored for multispectral images have been proposed. For instance, Salamati et al. proposed using both visible and NIR images as feature vector inputs to classifiers for categorizing materials within images. The relationship between visible and NIR information improved image-based material classification, resulting in more accurate material classification when NIR information was present [7]. Various convolutional neural network (CNN) algorithms, such as Alex-Net, Deep-CNN, and VGG have also been developed using RGB and NIR data as input layers for remote sensing image classification [8]–[11]. While these related studies demonstrate high classification accuracy, challenges remain, such as the need for large training samples in deep learning [12], making it crucial to achieve image classification with fewer training samples. Additionally, techniques for enhancing the color information in sensed images involve dividing the RGB color space into multiple color spaces and separating RGB color space images based on individual photodetectors to facilitate feature extraction. [2].

However, Jiang et al. demonstrated that image classification accuracy can be compromised when the subtle correlations between RGB and NIR images are not adequately accounted for in CNNs [13]]. Consequently, the enhancement of color information, which plays a pivotal role in image classification tasks, remains a significant challenge in the field of computer vision and image processing.

To address this challenge, we proposes a method that combines multispectral illumination capable of emitting visible and near-infrared light with a multispectral camera that can simultaneously capture visible and infrared regions. This approach enhance the information obtained through image measurement and further emphasizes color feature separation by analyzing the measured images for each photodetector. Furthermore, we adopt the Log-Linear Gaussian Mixture Network (LLGMN) [14] as a classifier, which can construct models based on statistical criteria from fewer samples without requiring numerous training samples like typical deep learning methods. This is expected to overcome the challenge of large training sample requirements in deep learning, potentially facilitating applications and social implementation across various fields.

In this paper, as a fundamental experiment for the pro-

[1]Taiga Eguchi, Wen Liang Yeoh, Hiroshi Okumura, and Osamu Fukuda are with Faculty of Science and Engineering, Saga University, 1 Honjomachi, Saga, 840-8502, Japan 23805101@cc.saga-u.ac.jp

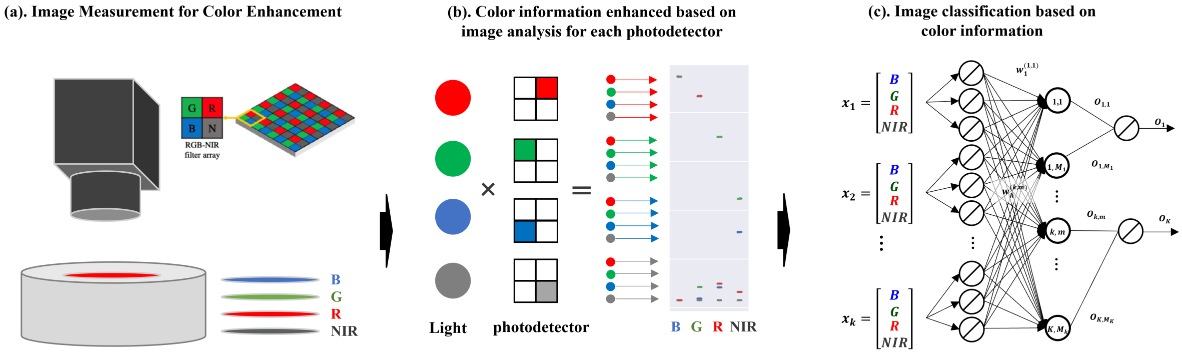

Fig. 1. Overview of the proposed method

posed color information enhancement technology, we focused on green tea, which possesses similar color characteristics, and verified the effectiveness of the proposed method, particularly emphasizing color information classification.

II. Methods

The overview of image classification utilizing enhanced color information arrays, as facilitated by the proposed multispectral lighting and multispectral camera system, is depicted in Fig. 1. The process consists of three main components: image acquisition for color enhancement, color information enhanced based on image analysis for each photodetector, and image classification based on color features. The details of these components are explained below.

A. System components

Fig. 2 illustrates the proposed measurement system. The dome-shaped illumination (MD2-100UPRLGB, SHIMATEC Y.K.) is controlled via a controller that regulates the intensity of the light. The multispectral camera (AD-080GE, JAI Inc.) is connected to a laptop computer (Let's note CF-SZ5, Panasonic) through a hub using a LAN cable. Image capture with the multispectral camera is performed using camera control software developed in the LabVIEW programming language. The sample, placed inside the dome-shaped illumination, is illuminated with R, G, B, and NIR light sources, and images are captured accordingly. The captured images consist of 24-bit RGB color images and 8-bit near-infrared (NIR) images. The image dimensions are 1024 $\times$ 248 pixels.

B. Extraction of enhanced color information

The image obtained from the measurement system undergoes color information enhancement as shown in Fig. 1-(b). First, the RGB color images captured under each illumination are separated into 8-bit color plane images for each photodetector. Subsequently, a 100×100 pixel region centered on each image is extracted from the three color plane images and the NIR image. The mean luminance is then calculated for these extracted images, which constitutes the enhanced color information. This enhanced color information is determined by the number of illumination spectral type used and the number of camera photodetector type. In this study, four types of illumination and four types of photodetectors are employed, resulting in a 16-dimensional enhanced color information array.

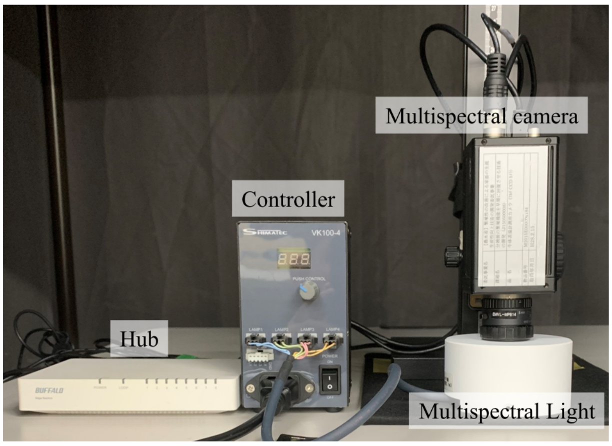

Fig. 2. System components

C. Discrimination using a Log-Linearized Gaussian Mixture Neural Network

Color information arrays are identified using LLGMN, which is a neural network with statistical structures [14].

The structure of LLGMN is shown in Fig. 1-(c). First, the feature vector $X \in R^d$ is pre-processed and transformed into the input vector $X \in R^d$. The first layer consists of H units corresponding to the dimensionality H of the input vector X and uses identity functions for the input-output functions of the units. The input-output relationship in the first layer is given by Equations (1) and (2), where ${}^{(1)}I_j$ is the input and ${}^{(1)}X_j$ is the output.

$$Y_{k,m} = \sum_{h=1}^{H} {}^{(1)}O_h w_h^{k,m} \tag{1}$$

$$^{(2)}O_{k,m} = \frac{\exp[Y_{k,m}]}{\sum_{\acute{k}=1}^{K} \sum_{\acute{m}=1}^{M_K} \exp[Y_{\acute{k},\acute{m}}]} \tag{2}$$

Note that $w^{K,M_K} = 0$.

The third layer consists of K event units and outputs the posterior probability of event $k(k = 1,\cdots,K)$. Unit k is coupled with M_k units $k,m(m = 1,\cdots,M_K)$ in the second layer. The input-output relationship is represented by Equations (3) and (4).

$$^{(3)}I_k = \sum_{m=1}^{M_K} {}^{(2)}O_{k,m} \tag{3}$$

$$O_k = ^{(3)} I_k \tag{4}$$

The LLGMN is trained using N sample data $x^{(n)}(n = 1,\cdots,N)$. Given N sample data points (training data), the log-likelihood function L is given by Equation (5).

$$L = \sum_{n=1}^{N} \sum_{k=1}^{K} T_k^{(n)} \log^{(3)} O_k \tag{5}$$

The output value of the network $^{(}3)O_k$ corresponds to the posterior probability $P(k \,|\, x^{(n)})$. For evaluation function J, we use Equation (6), which is Equation (5) with a negative sign, and learn to minimize it, that is, to maximize the likelihood.

$$J = \sum_{n=1}^{N} J_n = -\sum_{n=1}^{N} \sum_{k=1}^{K} T_k^{(n)} \log^{(3)} O_k \tag{6}$$

III. EXPERIMENTS

A. EXPERIMENTAL SETUP

To validate the proposed system, we utilized three green tea samples with similar colors that are challenging to distinguish with the human eye, as illustrated in Fig. 3. The measurements were conducted with the samples placed in transparent plastic cases. Each green tea sample was measured using a volume of 5 ml. The intensity of the multispectral illumination was determined by performing image measurements on the transparent plastic cases while adjusting parameters. Based on the results shown in Fig. 4, the illumination intensity for each light source was set to achieve approximately 90% of the maximum value (255) in the corresponding image brightness detected by the photodetector. The determined illumination control parameters were B: 15, G: 95, R: 75, and NIR: 65.

The experiments were conducted in a darkroom to ensure a stable environment. The illuminance within the darkroom was approximately 5-10 lx. For color classification, images were captured for each of the three classes and for each illumination spectrum, with 30 learning samples and 30 test samples per class. Consequently, there were 120 learning samples and 120 test samples per class. In the experimental procedure, we first confirmed the separation of the enhanced color information using scatter plots of the learning samples. Subsequently, we investigated the color classification accuracy and processing time of the LLGMN (Log-Linearized Gaussian Mixture Network) using the enhanced color information.

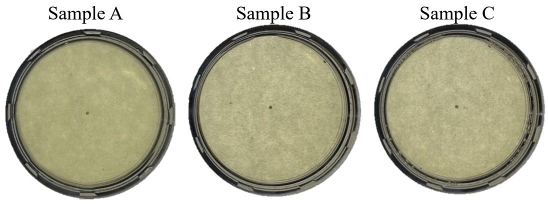

Fig. 3. Samples of three different green teas of similar color

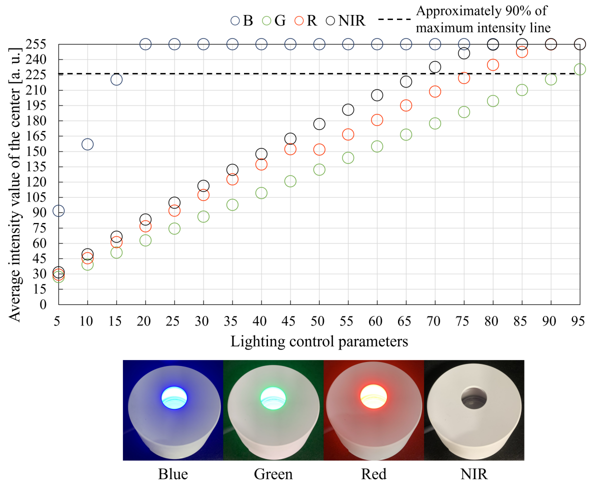

Fig. 4. Samples of three different green teas of similar color

B. EXPERIMENTAL RESULTS AND DISSCUSSION

1) Result of enhanced color information array: Fig. 5 shows examples of images acquired using the proposed measurement system. When the combination of illumination bands and photodetectors aligned, it was confirmed that the color information was emphasized in all three samples. However, upon visual comparison of any images from each sample, it remained difficult for the human eye to discern differences between samples. This is likely due to the original colors being indistinguishable at a level discernible by the human eye. Nonetheless, when focusing on and comparing the images illuminated with green light, which matches the sample color, subtle differences can be recognized.

Fig. 6 presents a comparison of the three samples' characteristics using the proposed method's color information array. Each section corresponds to a respective sample. The horizontal axis within each section represents the photodetectors, with the color features of average luminance for each irradiated spectral band plotted for each photodetector. This color feature plot represents the proposed 16-dimensional color information array. Comparing the distributions of the three samples reveals similar overall trends, which is interpreted as an indication of their similar color characteristics. However, for specific combinations of illumination and illumination color, subtle differences in reflection intensity among the three samples are accentuated, with these differences being particularly emphasized when blue and green illumination are applied. It is highly likely that this is due to the green color of the samples. It has long been known that

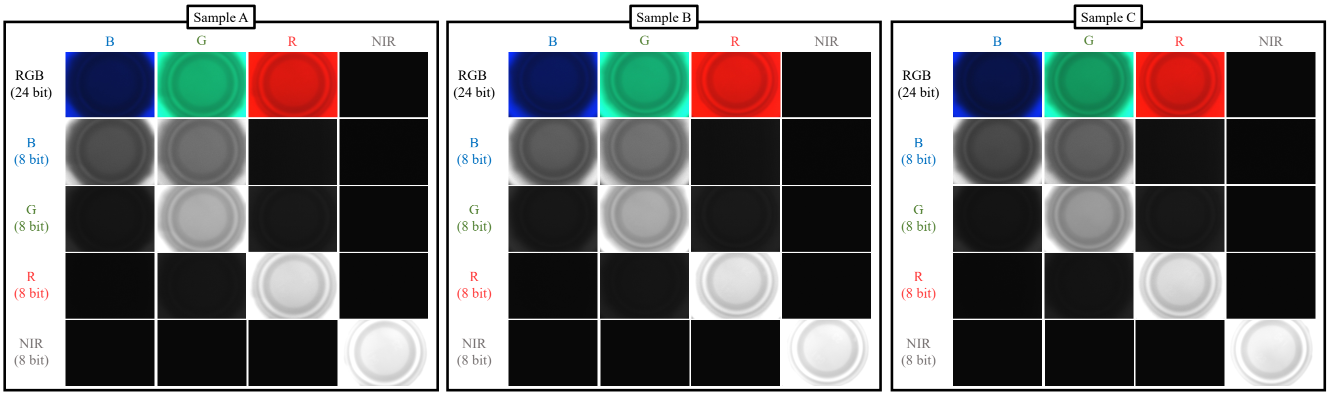

Fig. 5. Examples of images acquired by the proposed measurement system

illuminating a sample with light of a color similar to the sample's color emphasizes differences, and this effect appears to have manifested prominently here. Additionally, minor differences arise with other combinations of illumination and photodetectors, which are likely due to variations in reflection intensity caused by differences in the components contained within the liquids. The green tea samples used were created by dissolving powdered tea leaves in liquid. Differences in reflection intensity may have arisen due to factors such as the particle size of the powdered tea leaves, with these effects being reflected in the image luminance.

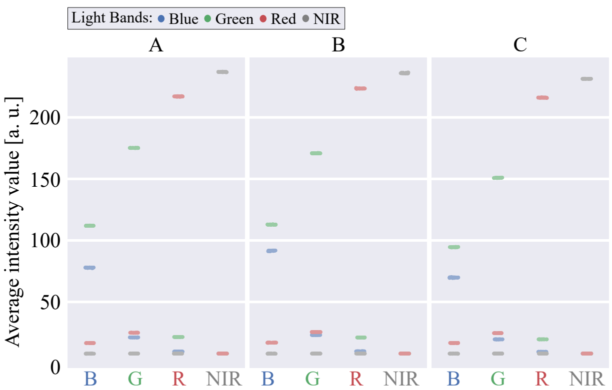

Fig. 6. Extraction results of the proposed color information array

2) Results of color classification using enhanced color information array: Fig. 7 shows the color classification results using LLGMN with the enhanced color information array as input, expressed as a confusion matrix. In Fig. 7, the vertical axis represents the actual sample names, while the horizontal axis shows the predicted sample names. The blue gradients indicate correct classifications, and red gradients represent misclassifications. The numerical values denote the proportions of correct and incorrect classifications, respectively. The experimental results demonstrate that even with only 120 image data per sample, the method was able to classify three types of green tea, which are difficult to distinguish by the human eye, with 100% accuracy. Furthermore, the time required to classify 90 samples, comprising 30 samples per class, was 0.71 seconds. These findings suggest that LLGMN, capable of high-speed processing despite the 16-dimensional input, may be well-suited for feature-based image classification.

This outcome indicates the effectiveness of the proposed color information enhancement technique. Moreover, they indicate the feasibility of color classification with a limited dataset compared to conventional deep learning approaches, experimentally demonstrating the significance of color information enhancement techniques in image classification.

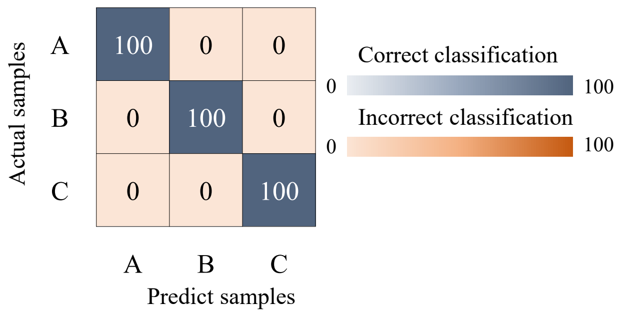

Fig. 7. Image classification result using LLGMN based on color information array

IV. CONCLUSIONS

In this study, we proposed a color classification method using LLGMN, which facilitates classification with minimal training samples. The method enhances color information, a crucial element in image classification, in both the visible (RGB) and near-infrared (NIR) regions using multispectral illumination and a multispectral camera. The effectiveness of the proposed method was evaluated using three types of green tea with similar color characteristics that are difficult for human eyes to distinguish. The experimental results showed that the proposed method successfully separated the color features by distinguishing differences in reflection intensity. Furthermore, by using LLGMN for feature-based image classification, it was demonstrated that 90 image samples could be classified with 100% accuracy in 0.71 seconds using only 120 training samples per class. These findings are expected to contribute to technological innovation in image classification across various fields where multispectral imaging is increasingly utilized, such as manufacturing, medicine, chemistry, and agriculture.

In the future, we plan to further validate the usefulness of the proposed method by conducting practical experiments that include not only color features but also shape features.

ACKNOWLEDGMENT

The authors would like to thank Jin Yien Lee of Saga University for his technical support and useful discussions during this research. The authors would also like to extend their gratitude to Shimatec Co., Ltd. for providing the demonstration equipment, which was instrumental in carrying out this research.

REFERENCES

[1] J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: https://doi.org/10.1109/CVPR.2009.5206848.

[2] S. N. Gowda and C. Yuan, " ColorNet: Investigating the Importance of Color Spaces for Image Classification, " Computer Vision – ACCV 2018, pp. 581–596, 2019, doi: https://doi.org/10.1007/978-3-030-20870-7_36.

[3] X. Soria, A. D. Sappa, and R. I. Hammoud, " Wide-Band Color Imagery Restoration for RGB-NIR Single Sensor Images, " Sensors, vol. 18, no. 7, p. 2059, Jul. 2018, doi: https://doi.org/10.3390/s18072059.

[4] X. Soria, A. D. Sappa and A. Akbarinia, "Multispectral single-sensor RGB-NIR imaging: New challenges and opportunities," 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), Montreal, QC, Canada, 2017, pp. 1-6, doi: https://doi.org/10.1109/IPTA.2017.8310105.

[5] J. Zhang, M. Li, Z. Sun, H. Liu, H. Sun and W. Yang, "Chlorophyll content detection of field maize using RGB-NIR camera", IFAC-PapersOnLine, vol. 51, no. 17, pp. 700-705, 2018, doi: https://doi.org/10.1016/j.ifacol.2018.08.114.

[6] G. Choe, S. -H. Kim, S. Im, J. -Y. Lee, S. G. Narasimhan and I. S. Kweon, "RANUS: RGB and NIR Urban Scene Dataset for Deep Scene Parsing," in IEEE Robotics and Automation Letters, vol. 3, no. 3, pp. 1808-1815, July 2018, doi: https://doi.org/10.1109/LRA.2018.2801390.

[7] N. Salamati, Clément Fredembach, and S. Süsstrunk, " Material Classification Using Color and NIR Images, " Final program and proceedings, vol. 17, no. 1, pp. 216–222, Jan. 2009, doi: https://doi.org/10.2352/cic.2009.17.1.art00040.

[8] H. Li et al., RSI-CB: A large scale remote sensing image classification benchmark via crowdsource data, 2017, [online] Available: https://arxiv.org/abs/1705.10450.

[9] X.-H. Han and Y.-W. Chen, "Generalized aggregation of sparse coded multi-spectra for satellite scene classification", ISPRS Int. J. Geo-Inf., vol. 6, no. 6, pp. 175, 2017, doi: https://doi.org/10.3390/ijgi6060175.

[10] V. Risojević, "Analysis of learned features for remote sensing image classification", Proc. 13th Symp. Neural Netw. Appl., pp. 1-6, Nov. 2016, doi: https://doi.org/10.1109/NEUREL.2016.7800145.

[11] Y. Zhong, F. Fei, Y. Liu, B. Zhao, H. Jiao and L. P. Zhang, "SatCNN: Satellite image dataset classification using agile convolutional neural networks", Remote Sens. Lett., vol. 8, no. 2, pp. 136-145, Feb. 2016, https://doi.org/10.1080/2150704X.2016.1235299.

[12] W. Rawat and Z. Wang, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review," in Neural Computation, vol. 29, no. 9, pp. 2352-2449, Sept. 2017, doi: https://doi.org/10.1162/neco_a_00990.

[13] J. Jiang, X. Feng, F. Liu, Y. Xu and H. Huang, "Multi-Spectral RGB-NIR Image Classification Using Double-Channel CNN," in IEEE Access, vol. 7, pp. 20607-20613, 2019, doi: https://doi.org/10.1109/ACCESS.2019.2896128.

[14] T. Tsuji, O. Fukuda, H. Ichinobe and M. Kaneko, "A log-linearized Gaussian mixture network and its application to EEG pattern classification," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 29, no. 1, pp. 60-72, Feb. 1999, doi: https://doi.org/10.1109/10.1109/5326.740670.