Multisensory Culinary Image Classification based on SqueezeNet and Support Vector Machine

1st Santirianingrum Soebandhi
Program Studi Manajemen
Universitas Wijaya Kusuma
Surabaya, Indonesia
santirianingrum@uwks.ac.id

2nd Adri Gabriel Sooai
Program Studi Ilmu Komputer
Pusat Unggulan Universitas bidang AI
untuk Pariwisata dan Agrikultur
Universitas Katolik Widya Mandira
Kupang, Indonesia
adrigabriel@unwira.ac.id

3rd Kristiningsih
Program Studi Manajemen
Universitas Wijaya Kusuma
Surabaya, Indonesia
kristiningsih@uwks.ac.id

4th Aryo Nugroho
Program Studi Sistem Informasi
Universitas Narotama
Surabaya, Indonesia
aryo.nugroho@narotama.ac.id

***Abstract*— Culinary tourism or food tourism can be a competitive advantage for Indonesian tourism. Indonesia with diverse cultures and ethnicities certainly has local culinary characteristics that are not found in other regions. Although culinary tourism research has been conducted, not many have examined culinary tourism from multisensory experience analysis through a big data approach to obtain datasets that can be processed to visualize tourist experiences. This study aims to classify the multisensory experienced by tourists when doing culinary tours utilizing deep learning SqueezeNet for image extraction and Support Vector Machine with Linear Kernel for image classification. Used primary dataset consisting of three image class, consisting of Vibes, Place and Food. This study succeeded in classifying multisensory culinary images with an accuracy rate of 98.6%.**

Keywords—multisensory, image classification, culinary tourism, support vector machine, squeezenet

I. INTRODUCTION

Culinary tourism can be a competitive advantage for the development of tourist destinations. Each region certainly has local food with different tastes, raw materials, processing and presentation methods. This can be a tourist attraction because it can provide a different sensation when eating local food. With enormous geographical and cultural diversity, Indonesian cuisine is certainly rich in variety and taste. The diversity of traditional dishes is a challenge for the government to promote to the international market [1]. Culinary connoisseurs have diverse inputs from their respective culinary tours. This is as a result of what is recorded by the human senses such as sight, hearing, food taste, smell, touch. The sense of sight can more or less represent or conjecture the initial value of some other Senses [2], [3]. The results of human vision that summarize several senses or in other words Multisensory can be compiled from several groups of culinary images ranging from photos of food, atmosphere and location of culinary centers. Several studies around culinary tourism that collect tourist feelings and responses in the form of tweets and a number of food photos from online social media have been conducted [4]. Other studies only rely on a variety of inputs from investors who first have culinary tourism experience from thematic locations [5]. The relationship between culinary tourism photos (food: food and beverage, vibes: atmosphere, and place: coziness) all represent the human senses (sight, taste, and touch) is closely related because culinary tourism not only includes the sensation of food, but also offers a rich multisensory experience [6].

A. Sighting and Photos of Culinary Tours

Multimedia technology has an important role in the development, promotion, and presentation of cultural tourism[7], [8] and especially culinary. The linkage between multimedia technology and cultural and culinary tourism involves the use of various multimedia elements to enhance the tourist experience and promote the cultural and culinary heritage of a region. In a simple form, a collection of photos can represent places, activities, and various things about culinary tourism. Photos of culinary tourism play an important role in attracting human attention. Vision is the first sense used when a person looks at photos of food, atmosphere, or culinary attractions. Appetizing images, attractive food layouts, charming place décor, or beautiful natural scenery will influence interest and desire to taste food and visit the place [9].

B. Taste and Food

Taste is the main sense involved when a person tastes food. When people look at photos of food, the taste is almost carried away by imagining the culinary sensation they will get. Taste is a key element in the overall culinary experience, and food photos can stimulate one's taste buds and increase the desire to try the dish [10].

C. Atmosphere and Touch Experience

The atmosphere and tactile experience includes a variety of elements, including room décor, furniture, and food textures. Photos of the atmosphere in culinary attractions can present different nuances, such as relaxed, romantic, or adventurous. In addition, food photos can also show the texture of a dish, for example, food that is crispy, soft, or chewy. This tactile experience can affect a person's perception of the place and make them feel more connected to their surroundings [11].

This study intends to contribute to a culinary image classifier model with relatively high accuracy, and represents several multisensory related culinary tourism. The dataset used is a primary dataset divided into four classes, namely Vibes, Coziness, Place and Food. With a total of 119 images that go through a VGG-16 assisted vitur extraction process for further processing using the Support Vector Machine with a Linear kernel. The acquisition tools used in this study were sourced from various types of gadgets owned by surveyors with an average image resolution of 4080 x 1780 pixels, density of 72 pixels, bit depth of 24, RGB color, exposure 1/100, ISO speed 179, maximum aperture 1.69. Processing using 1GHz CPU device, 4 GB RAM, without GPU.

The systematics of reporting this research begins with a presentation that describes the latest trends around culinary tourism and how tourists interact with their culinary locations and attractions. Furthermore, various related studies were presented which became the basis for the contribution of this research. Then described how this research was carried out through a sequence of stages. A review of the results that have been achieved in this research is also presented and closed with conclusions and suggestions for further research.

II. Related Works

A study on the introduction and classification of food has been conducted using a deep learning classifier. The research resulted in a food image recognition method called NutriNet. This deep learning architecture is used to recognize the image of real food as well as the image of fake food and drinks. This architecture is a modification of alexnet by increasing the pixel size from 256x256 to 512x 512 for each image processed. Modifications were also made to the alexnet by adding a convolution layer. 631 food images were used which were divided into 55 food classes. In addition, processors other than NutriNet are used, classifiers using Fully Convolutional Network 8s [12]. However, the study only produced accuracy at 92.18%.

Still within the scope of food image recognition, research has been conducted to classify food in several groups, namely unprocessed, processed and ultra processed and other culinary. The study used secondary datasets, namely EgocentricFood and UECFOOD, then a primary dataset containing food images was taken using a makeshift gadget camera and labeling was carried out on each group of images [13]. However, the study only succeeded in achieving an accuracy rate of 90% classification for mAP and 86% for NOVA food Classifier

Food classification using machine learning technology is also carried out for foods that are considered to have different nutritional values. The EfficientNetB0 algorithm is used in transfer learning mode. A total of 101 types of food in traditional and foreign food groups were classified. The experimental process goes through several stages, namely feature transfer, fine tuning, pre-trained model and result acquisition. As a benchmark, several comparison algorithms are also used, namely CNN, Random Forest, InceptionV3 and GoogleNet [14]. However, the use of EfficientNetB0 used in this study was only able to produce a classifier accuracy rate of 80% and is equivalent to its comparator GoogleNet, while other classifiers did not exceed 80%.

Similar research in culinary matters is devoted to sorting out foods for people who avoid obesity and how to choose healthy foods has been done. The dataset they used was an available secondary dataset under the name Food-101. Transfer learning and adjustment of several classification algorithms such as InceptionV3 and V4 are used to recognize food in the form of digital images. Several stages were carried out in the experimental process, namely augmentation, multi-crop and initial processing of other food images [15]. However, this study only succeeded in producing a classifier accuracy of the dataset used, which was 85%. The overall related research that has been carried out to produce food image modeling can be seen in Table 1

Table 1 Related Research

Author	Method	Accuracy
Mezgec, 2019	Nutrinet & Fully Convolutional Network 8s	92.18%
Elbassuoni, 2022	mAP & NOVA food Classifier	86% & 90%
Vutkur, 2022	EfficientNetB0	80%
Shen, 2020	Inception V3 & V4	85%

III. Method

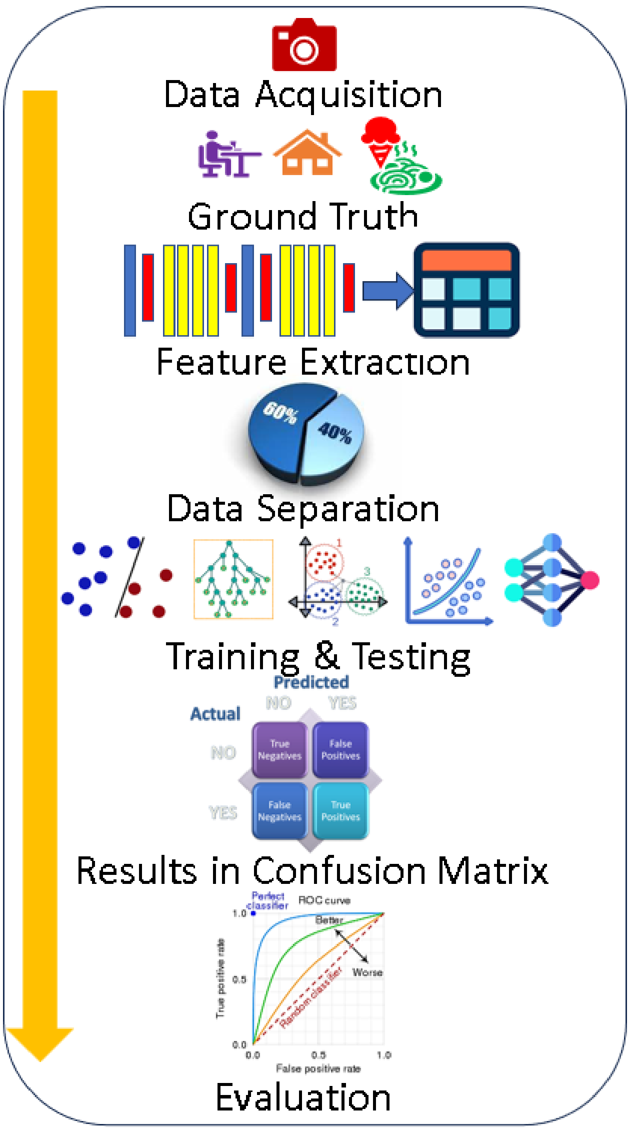

Figure 1 Experiment Flow

This research was conducted in several stages starting with image acquisition to create a dataset of culinary tourism visits to various places of interest. Followed by ground truth, feature extraction, separation of training and testing data, modeling, results and obtaining evaluation, the entire process of this experiment can be seen in Figure 1

A. Dataset Compilation

The grouped images are arranged in three multisensory classes: those representing Feelings are labeled "Vibes", then the results of vision of culinary products are labeled "Food" and the results for Hearing and Touch sensors are labeled "Place". After all images have been successfully acquired and grouped under predetermined labels, the next step is to do Ground Truth by selecting each image in more detail by removing photos between similar classes because it will decrease accuracy. It is important to obtain a high-precision training model.

B. Feature extraction

The next stage is the feature extraction process using pre-trained Deep Learning, namely SqueezNet [16]–[18]. In this process, all 119 Photos are converted into numeric data in the form of Rows and Columns with a total of 119 rows and 1000 columns. Each column contains vector features of each image as a result of extraction performed using the deep model, SqueezeNet;. The feature extraction process using SqueezeNet can be followed on pseudocode as shown in Figure 2. The extracted data is then divided in half by a 60:40 combination for training and testing. The purpose of dividing data into two groups, training and testing, is to ensure that the model that has been created can be used to test data outside the model in the hope of producing identification that has a high level of accuracy.

C. Modeling

Some classifiers used as comparisons in experiments that have been carried out are Support Vector Machine (SVM) [19], Neural Network (NN), kNN, Random Forest (RF), Logistic Regression (LR) and Naïve Bayes (NB). The first modeling process or classification performed using is done using a Linear kernel and 3-Fold Cross Validation. 3-Fold is selected after going through iterations that show the highest accuracy results. Several other classifier will also be tested against the dataset for comparison, thus SVM with a linear kernel is certain to outperform other traditional classifiers for this multisensory primary dataset. The use of SVM with a linear kernel was chosen as the highest accuracy. In general, SVM will use the train dataset in notation$(x_1, y_1), \ldots, (x_n, y_n)$ where the class is represented by y_i which consists of x_i. Grouping using a linear hyperplane as shown in the formula (1). Parameter w denoting a vector against a hyperplane, while b is offset.

$$wTx - b = 0 \tag{1}$$

$$\emptyset(z) = \frac{1}{1 + e^{-1}} \tag{2}$$

$$Q(w) = \frac{1}{n}\sum_{i=1}^{n} Q_i(w) \tag{3}$$

```
import numpy as np
from Orange.data import Domain, Table
from Orange.widgets.utils.signals import Output
data = Table("..\dataMultisensory")
def extract_squeezenet_embedding(image_path):
embedding = np.random.rand(512)
return embedding
num_channels = 512
attributes = [Orange.data.ContinuousVariable(f"channel_{i}") for i in range(num_channels)]
embedding_domain = Domain(attributes)
embedded_data = []
for instance in data:
embedding= extract_squeezenet_embedding(instance["image_path"])
embedded_data.append(embedding)
embedding_table = Table(embedding_domain, embedded_data)
output_data = embedding_table
Output(output_data)
```

Figure 2 Feature Extraction process using Squeezenet

Next is the Neural Network, also known as the Multilayer Perceptron with backward propagation prepared as a classifier for performance comparison. Assisted logistic activation function according to the formula (2) and Stochastic Gradient Descent according to the formula (3) in the form of stochastic estimates. NN works with a number of neurons in a number of hidden layers. $\emptyset$ is logistic function notation, z is horizontal Asymptote Notation, Q_i is the sum result of each iteration i, while w is a parameter to minimize Q.

kNN as a guided non-parametric classifier, kNN produces an output i.e. class membership. Each observation is classified by proximity between objects and represented in positive k values using a range of small numbers such as 1,3,5,10 and 20 [20], [21]. This algorithm can utilize a variety of distance metrics for the classification process, including Mahalanobis, Chebyshev, Manhattan, Minkowski, Hamming and Euclidean. The proximity between objects in the observation is calculated using the Euclidean distance metric in the formula (4).

$$d(e, f) = \sum_{i=1}^{n} (e_i - f_i)^2 \tag{4}$$

d represents the distance between two objects represented by the notation e and f

Naïve Bayes uses a probability-based classifier, NB is also included and can be used for image processing. Data modeling in NB processes observations that have been grouped and arranged in classes [22], [23]. Usually grouping is done through the stages of ground truth. Two NB models are available, namely simple bayes network and kernel density estimation. The simple bayes network used in this study can be seen in the formula (5).

$$p(C_k|x) = \frac{p(x|C_k)p(C_k)}{p(x)} \tag{5}$$

The next comparison classifier is Random Forest. Random forest or random decision forest is a classification method that utilizes ensemble learning, which is a prediction method that uses several stages of learning. One of the ensemble learning algorithms used in random forests is bootstrap aggregation, otherwise known as bagging [24], [25]. Random forest also use regression methods and other tasks through the preparation of many decision trees, with formulas such as in (6).

$$\hat{f} = \frac{1}{B}\sum_{b=1}^{B} f_b(x') \quad (6)$$

The training data is denoted as $X = x_1, \dots, x_n$, while the responses are notated with $Y = y_1, \dots, y_n$, and repetition bagging is notated with B iteration. The amount of training data is denoted as n, samples with content replacement are notated with X_b, Y_b regression trees are denoted as f_b on X_b, Y_b, and after the training process, the prediction is notated as x'.

The final stage is the evaluation process carried out by displaying the results of classification of datasets using a confusion matrix. Confusion matrix plays an important role in explaining classification and prediction results by presenting information about actual prediction results and prediction results made by the model [25], [26].

D. Measuring Classifier Performance

Advanced evaluation is to compare the performance of each classifier using several other quantities Using information from the confusion matrix, it can calculate several important evaluation metrics, Some quantities used to determine classifier performance are Area Under the Receiver Operating Characteristic Curve (AUC-ROC) or abbreviated AUC [26], Classification Accuracy(CA), F1, Precision and Recall. All of these variables work based on several actual prediction and observation results, namely, results where the model correctly predicts a positive class that is True Positive (TP), results where the model predicts a negative class correctly that is True Negative (TN), results where the model incorrectly predicts a false positive class that is False Positive (FP), and results where the model incorrectly predicts a negative class, i.e. False Negative(FN) [27]. From these four quantities form several formulas, namely::

$$CA = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$Recall(True\ Positive\ Rate) = \frac{TP}{TP + FN} \quad (8)$$

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$F1 = 2 * \frac{Recall * Precision}{Recall + Precision} \quad (10)$$

$$False\ Positive\ Rate = \frac{FP}{FP + TN} \quad (11)$$

The AUC value is obtained by mapping the formula (8) and (11) to form a Receiver Operator Characteristic Curve (ROC), where the highest or best value is close to 1.0 which means the best measurement is created.

IV. Results and Discussion

The overall image obtained from the data acquisition is divided into three classes, namely Vibes represent the atmosphere in the culinary center, Place represents information about the place, can be in the form of restaurants or cafes or street vendors and Food represents culinary diversity encountered by tourists. The illustration can be seen in Figure 3. This research produces several outputs that can be used as a reference that a high accuracy model in accordance with the objectives of this research is achieved. The output is confusion matrix, accuracy between classifiers and comparison of models based on ROC.

The first output, the confusion matrix table, displays the results of predictions that have been made using modeling of datasets with three classes, namely Food, Place and Vibes. Based on 60:40 data separation, a total of 72 images trained were tested using 40% of the test data. The result was that 28 Food labels were identified, while one image was incorrectly identified as the Place label. For the other two labels, Vibes and Place, they were all identified 100%. Everything can be seen in Table *2*

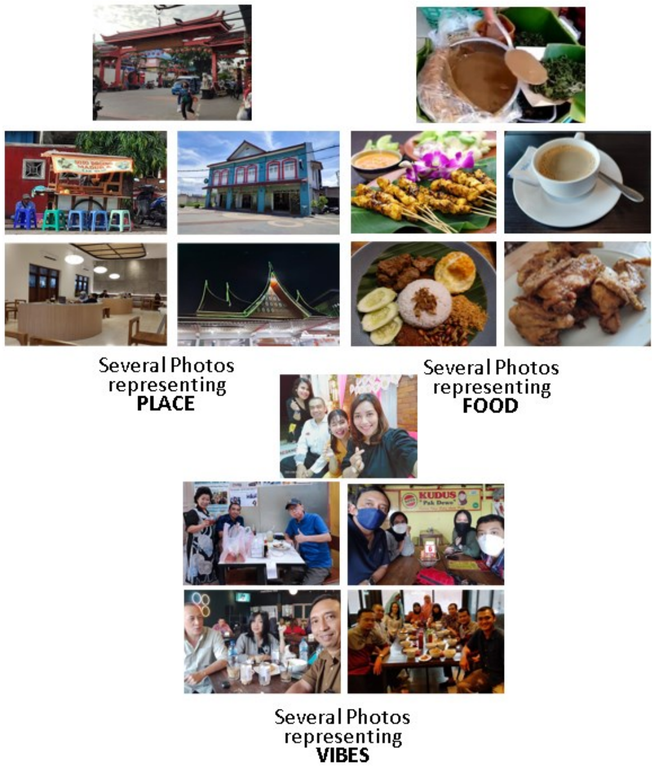

Figure 3 Class representation

Table 2 Confusion Matrix of SVM Linear tested **using 40% data.**

		Predicted			
		Food	Place	Vibes	∑
Actual	Food	28	1	0	29
	Place	0	30	0	30
	Vibes	0	0	13	13
	∑	28	31	13	72

As can be seen on Table 3, by obtaining the prediction value displayed in the confusion matrix, the second output that becomes a benchmark for the success of experiments that have been carried out is the accuracy table between classifications. In this table some of the quantities used are Area Under Curve, Classification Accuracy, F1, Precision and Recall.

Sequentially from the classifier with the highest accuracy number, namely SVM, in order to the lowest classifier, namely Tree. Broadly speaking, there are three groups of classification accuracy quantities, namely the number group above 94% to 98%, the 80% group and the 70% group. Although with a linear kernel for SVM classifiers, it turns out that the extraction results in the form of image datasets that have been converted into numerical data can be selected on the support vector machine hyperplane to produce models with accuracy that is good enough to beat previous studies.

TABLE 3 ACCURACY BETWEEN CLASSIFIERS

Model	AUC	CA	F1	Precision	Recall
SVM	99%	98.6%	98.6%	98.7%	98.6%
LR	99.4%	97.2%	97.2%	97.4%	97.2%
kNN	96.2%	95.8%	95.8%	92.6%	95.8%
NN	98.8%	94.4%	94.4%	94.7%	95.8%
NB	97.6%	84.7%	85.3%	90.2%	94.4%
RF	95.1%	84.7%	83.8%	83.7%	84.7%
Tree	84.8%	73.6%	74.4%	76.7%	73.6%

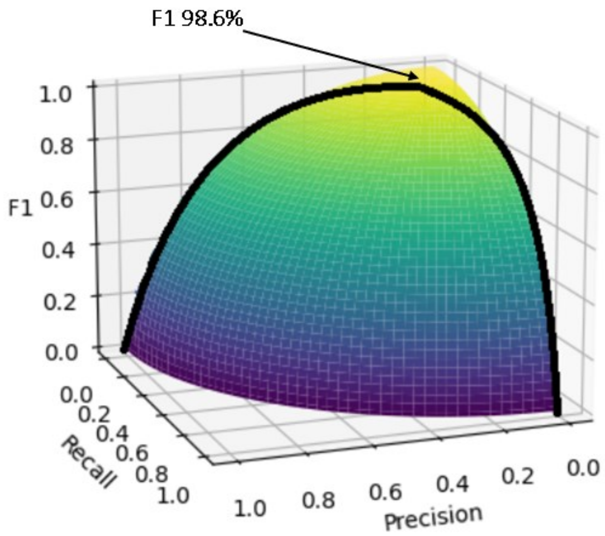

Figure 4 Visualization of the Highest F1 score

Referring to the accuracy between each classifier, for that the comparative evaluation that can be used as a measure for accuracy is the value of F1. The following is shown a visualization of the F1 position of the SVM classifier with a Linear Kernel that managed to reach 98.6% depicted on the back of the spherical quarter graph as shown in Figure 4

TABLE 4 MODEL COMPARISON BASED ON ROC

	SVM	LR	kNN	NN	NB	RF	Tree
SVM		0.296	0.819	0.737	0.973	0.765	0.997
LR	0.704		0.817	0.733	0.966	0.722	0.997
kNN	0.181	0.183		0.186	0.337	0.569	0.979
NN	0.263	0.267	0.814		0.848	0.715	1.000
NB	0.027	0.034	0.663	0.152		0.685	0.998
RF	0.235	0.288	0.431	0.285	0.315		0.879
Tree	0.003	0.003	0.021	0.000	0.002	0.121	

Based on the previous two outputs then the last one output to ensure classifier performance is displayed in the comparison table of each model based on the ROC, TABLE 4 shows the probability that the score for the models in the row is higher than the model in the column, a small number indicates the probability that the difference is negligible. A comparative analysis of classifiers based on ROC can be read on TABLE 4 represented by reading the first line i.e. the probability of SVM to Logistic Regression is greater by numbers 0.296, notated with p(SVM>LR)=0.296, further against k-Nearest Neighbor denoted as p(SVM>kNN)=0.819, then against Neural Network as p(SVM>NN)=0.737, and probability against Naïve Bayes as p(SVM>NB)=0.973, followed probability against Random Fores as p(SVM>RF)=0.765, and finally is the probability of the Decision Tree as p(SVM>Tree)=0.997.

Why use SqueezeNet and not others, it is based on a smaller number of parameters than other feature extraction algorithms. In comparison InceptionV3 produces 2048 vector features, VGG16 produces 4096 vector features while SqueezeNet only generates 1000 vector features containing information from the image. Thus, this quantity cuts down the iteration process in the training and testing stages of the data. Of the several classifiers used, namely SVM, NN, kNN, RF, LR, Tree and NB, the highest classifier accuracy is obtained when using SVM. The combination of SqueezeNet and SVM for the feature extraction and classification process in this study is the choice that produces the model with the highest accuracy based on the dataset used. Then, all the results of classification using various algorithms can be seen in TABLE 3. Furthermore, all of these results are seen in probability using TABLE 4, namely ROC, the relationship between TABLE 3 and TABLE 4 is complementary, although it is easier to read the results in TABLE 3, however ROC is completeness to justify results in machine learning.

V. CONCLUSION

Some of the human senses represented by the three photo classes, such as Vibes, Food and Place, include sight, taste and atmosphere, and touch, interact and influence each other in a culinary tourism experience. For example, an appetizing image of food can trigger hunger and a desire to taste the dish. Likewise, a comfortable atmosphere and attractive views can increase satisfaction when eating. The holistic experience creates lasting and enjoyable memories for visitors. Accordingly, culinary tourism photos have an important role in influencing human perception and experience of food, atmosphere, and places. Sight provides the initial understanding of what is on offer, taste is the final judgment on food, and atmosphere and touch contribute to an overall experience that engages all the human senses. The paper studied a classification system for the culinary multisensory problem. The study targeted to classify the multisensory experienced by tourists when doing culinary tours by considering SqueezeNet deep learning for image extraction and Support Vector Machine with Linear Kernel for image classification. The proposed image extractor and classifier resulted in very high accuracy for the multisensory culinary image classification problem.

ACKNOWLEDGMENT

This research is funded by DRTPM DIKTI fundamental research scheme 2023, with contract number 183/E5/PG.02.00.PL/2023, dated June 19, 2023 between DRTPM and LLDIKTI VII, 005/SP2H/PT/LL7/2023 dated June 19, 2023 between LLDIKTI and UWKS,

04/LPPM/UWKS/VI/2023 dated June 21, 2023 between LPPM-UWKS and Researchers.

REFERENCES

[1] S. Wijaya, "Indonesian food culture mapping: a starter contribution to promote Indonesian culinary tourism," *Journal of Ethnic Foods*, vol. 6, no. 1, p. 9, Sep. 2019, doi: 10.1186/s42779-019-0009-3.

[2] C. Spence, K. Motoki, and O. Petit, "Factors influencing the visual deliciousness / eye-appeal of food," *Food Quality and Preference*, vol. 102, p. 104672, Dec. 2022, doi: 10.1016/j.foodqual.2022.104672.

[3] S. E. Kemp, I. Nyambayo, L. Rogesr, T. Sanderson, and C. B. Villarino, "Trends in food sensory science," *Food Science and Technology*, vol. 35, no. 4, pp. 46–50, 2021, doi: 10.1002/fsat.3504_13.x.

[4] S. Jiang and A. Ngien, "The Effects of Instagram Use, Social Comparison, and Self-Esteem on Social Anxiety: A Survey Study in Singapore," *Social Media + Society*, vol. 6, no. 2, p. 2056305120912488, Apr. 2020, doi: 10.1177/2056305120912488.

[5] L. Xie-Carson, P. Benckendorff, and K. Hughes, "Keep it #Unreal: Exploring Instagram Users' Engagement With Virtual Influencers in Tourism Contexts User engagement with virtual influencers," *Journal of Hospitality & Tourism Research*, p. 10963480231180940, Jul. 2023, doi: 10.1177/10963480231180940.

[6] J. Youssef and C. Spence, "Náttúra by Kitchen Theory: An immersive multisensory dining concept," *International Journal of Gastronomy and Food Science*, vol. 24, p. 100354, Jul. 2021, doi: 10.1016/j.ijgfs.2021.100354.

[7] I. G. A. Dharmayasa, S. Sumpeno, I. K. E. Purnama, and A. G. Sooai, "Exploration of prayer tools in 3D virtual museum using leap motion for hand motion sensor," in *2017 TRON Symposium (TRONSHOW)*, Dec. 2017, pp. 1–8. doi: 10.23919/TRONSHOW.2017.8275079.

[8] A. G. Sooai, A. Nugroho, M. N. A. Azam, S. Sumpeno, and M. H. Purnomo, "Virtual artifact: Enhancing museum exhibit using 3D virtual reality," in *2017 TRON Symposium (TRONSHOW)*, Dec. 2017, pp. 1–5. doi: 10.23919/TRONSHOW.2017.8275078.

[9] O. E. S. Rodriguez, S. Masuko, and T. Yamanaka, "The Impression of Deliciousness through Food Photography," *International Journal of Affective Engineering*, vol. 21, no. 1, pp. 67–76, 2022, doi: 10.5057/ijae.TJSKE-D-20-00076.

[10] M. Paakki, M. Sandell, and A. Hopia, "Visual attractiveness depends on colorfulness and color contrasts in mixed salads," *Food Quality and Preference*, vol. 76, pp. 81–90, Sep. 2019, doi: 10.1016/j.foodqual.2019.04.004.

[11] J. E. Machin, E. Moscato, and C. Dadzie, "Visualizing food: photography as a design thinking tool to generate innovative food experiences that improve food well-being," *European Journal of Marketing*, vol. 55, no. 9, pp. 2515–2537, Jan. 2021, doi: 10.1108/EJM-02-2020-0141.

[12] S. Mezgec and B. K. Seljak, "Using Deep Learning for Food and Beverage Image Recognition," in *2019 IEEE International Conference on Big Data (Big Data)*, Los Angeles, CA, USA: IEEE, Dec. 2019, pp. 5149–5151. doi: 10.1109/BigData47090.2019.9006181.

[13] S. Elbassuoni *et al.*, "DeepNOVA: A Deep Learning NOVA Classifier for Food Images," *IEEE Access*, vol. 10, pp. 128523–128535, 2022, doi: 10.1109/ACCESS.2022.3227769.

[14] V. G., P. Vutkur, and V. P., "Food classification using transfer learning technique," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 225–229, Jun. 2022, doi: 10.1016/j.gltp.2022.03.027.

[15] Z. Shen, A. Shehzad, S. Chen, H. Sun, and J. Liu, "Machine Learning Based Approach on Food Recognition and Nutrition Estimation," *Procedia Computer Science*, vol. 174, pp. 448–453, Jan. 2020, doi: 10.1016/j.procs.2020.06.113.

[16] A. Gholami *et al.*, "SqueezeNext: Hardware-Aware Neural Network Design." arXiv, Aug. 27, 2018. doi: 10.48550/arXiv.1803.10615.

[17] C. Y. Jerandu *et al.*, "Image Classification of Decapterus Macarellus Using Ridge Regression," in *2022 8th International Conference on Education and Technology (ICET)*, Oct. 2022, pp. 81–86. doi: 10.1109/ICET56879.2022.9990820.

[18] A. J. Lado *et al.*, "Comparison of Neural Network and Random Forest Classifier Performance on Dragon Fruit Disease," in *2021 International Electronics Symposium (IES)*, Sep. 2021, pp. 287–291. doi: 10.1109/IES53407.2021.9593992.

[19] S. S. Teja Gontumukkala, Y. S. Varun Godavarthi, B. R. Ravi Teja Gonugunta, R. Subramani, and K. Murali, "Analysis of Image Classification using SVM," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Jul. 2021, pp. 01–06. doi: 10.1109/ICCCNT51525.2021.9579803.

[20] A. Nugroho, S. Sumpeno, and M. H. Purnomo, "A Decision Guidance for Solving Success Rate Political Campaign Using Distance Weighted kNN in Nassi-Shneiderman Framework," *IJIES*, vol. 14, no. 2, pp. 410–420, Apr. 2021, doi: 10.22266/ijies2021.0430.37.

[21] A. G. Sooai *et al.*, "Comparison of Recognition Accuracy on Dynamic Hand Gesture Using Feature Selection," in *2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM)*, Nov. 2018, pp. 270–274. doi: 10.1109/CENIM.2018.8711397.

[22] S. Márquez-Sánchez, I. Campero-Jurado, J. Quintanar-Gómez, S. Rodríguez, and J. M. Corchado, "Smart Belt Design by Naïve Bayes Classifier for Standard Industrial Protection Equipment Integration," *International Journal of Artificial Intelligence™*, vol. 18, no. 2, Art. no. 2, Feb. 2020.

[23] K. A. Rahman *et al.*, "Crime Rate Mapping in Bandung City Area with Online Media Data by Using Naïve Bayes Method," *IIJ*, vol. 14, no. 1, pp. 21–26, 2022.

[24] N. Z. Fanani *et al.*, "Two Stages Outlier Removal as Pre-processing Digitizer Data on Fine Motor Skills (FMS) Classification Using Covariance Estimator and Isolation Forest," *IJIES*, vol. 14, no. 4, pp. 571–582, Aug. 2021, doi: 10.22266/ijies2021.0831.50.

[25] A. Kulkarni, D. Chong, and F. A. Batarseh, "5 - Foundations of data imbalance and solutions for a data democracy," in *Data Democracy*, F. A. Batarseh and R. Yang, Eds., Academic Press, 2020, pp. 83–106. doi: 10.1016/B978-0-12-818366-3.00005-8.

[26] O. A. Montesinos López, A. Montesinos López, and J. Crossa, *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. Cham: Springer International Publishing, 2022. doi: 10.1007/978-3-030-89010-0.

[27] - Istiadi *et al.*, "Classification of Tempeh Maturity Using Decision Tree and Three Texture Features," *JOIV : International Journal on Informatics Visualization*, vol. 6, no. 4, Art. no. 4, Dec. 2022, doi: 10.30630/joiv.6.4.983.

LETTER OF ACCEPTANCE

2023 IEEE 9th Information Technology International Seminar (ITIS)

Paper ID : 1570935475

Paper Title : Multisensory Culinary Image Classification Based on SqueezeNet and Support Vector Machine

Dear Santirianingrum Soebandhi, Adri Sooai, Kristiningsih, Aryo Nugroho

On behalf of the Organizing Committee of 2023 IEEE 9th Information Technology International Seminar (ITIS), I am pleased to inform you that based on the recommendations of the reviewers and the Technical Program Committees, your paper identified above has been ACCEPTED for oral presentation.

ITIS conference papers were reviewed by international experts. The acceptance ratio is controlled below 65%. The paper will be submitted after being revised according to the reviewer's suggestions in the proceedings of 2023 IEEE 9th Information Technology International Seminar (ITIS) by IEEE Xplore after registration and presentation.

We look forward to welcoming you virtually or in person in Batu – Malang, Indonesia, on October 18 - 20, 2023.

Yours Sincerely,

INFORMATION TECHNOLOGY INTERNATIONAL SEMINAR

Dr. Rr. Ani Dijah Rahajoe, ST., M.Cs.
General Chair
2023 IEEE 9th Information Technology International Seminar (ITIS)

CERTIFICATE OF APPRECIATION

No. 25/UN 63.7/ITIS/2023

Santirianingrum Soebandhi, Adri Sooai, Kristiningsih, Aryo Nugroho

"Multisensory Culinary Image Classification Based on SqueezeNet and Support Vector Machine"

As "Author" at

The 2023 IEEE 9th Information Technology International Seminar (ITIS)

October, 18th - 20th 2023

Batu Malang, Indonesia

ITIS INFORMATION TECHNOLOGY INTERNATIONAL SEMINAR

Dr. Rr. Ani Dijah Rahajoe, S.T, M.Cs

Conference General Chair

CERTIFICATE OF APPRECIATION

No. 25/UN 63.7/ITIS/2023

Santirianingrum S & Adri Gabriel Sooai

As "Presenter" at

The 2023 IEEE 9th Information Technology International Seminar (ITIS)

October, 18th - 20th 2023

Batu Malang, Indonesia

Dr. Rr. Ani Dijah Rahajoe, S.T, M.Cs

Conference General Chair