Hindawi
Evidence-Based Complementary and Alternative Medicine
Volume 2022, Article ID 7684714, 14 pages
https://doi.org/10.1155/2022/7684714

Research Article

Research of the Distribution of Tongue Features of Diabetic Population Based on Unsupervised Learning Technology

Jun Li,[1] **Longtao Cui**,[1] **Liping Tu**,[1] **Xiaojuan Hu**,[2] **Sihan Wang**,[1] **Yulin Shi**,[1] **Jiayi Liu**,[1] **Changle Zhou**,[3] **Yongzhi Li**,[4] **Jingbin Huang**,[1] **and Jiatuo Xu**[1]

[1]*School of Basic Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China*
[2]*Shanghai Collaborative Innovation Center of Health Service in Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China*
[3]*Department of Intelligent Science and Technology, Xiamen University, Xiamen, Fujian, China*
[4]*China Astronaut Research and Training Center, Beijing, China*

Correspondence should be addressed to Jingbin Huang; jmy61@163.com and Jiatuo Xu; xjt@fudan.edu.cn

Received 22 November 2021; Revised 4 June 2022; Accepted 11 June 2022; Published 5 July 2022

Academic Editor: Youhua Wang

Copyright © 2022 Jun Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. The prevalence of diabetes increases year by year, posing a severe threat to human health. Current treatments are difficult to prevent the progression of diabetes and its complications. It is imperative to carry out individualized treatment of diabetes, but current diagnostic methods are difficult to specify an individualized treatment plan. *Objective*. Clarify the distribution law of tongue features of the diabetic population, and provide the diagnostic basis for individualized treatment of traditional Chinese medicine (TCM) in the treatment of diabetes. *Methods*. We use the TFDA-1 tongue diagnosis instrument to collect tongue images of people with diabetes and accurately calculate the color features, texture features, and tongue coating ratio features through the Tongue Diagnosis Analysis System (TDAS). Then, we used K-means and Self-organizing Maps (SOM) networks to analyze the distribution of tongue features in diabetic people. Statistical analysis of TDAS features was used to identify differences between clusters. *Results*. The silhouette coefficient of the K-means clustering result is 0.194, and the silhouette coefficient of the SOM clustering result is 0.127. SOM Cluster 3 and Cluster 4 are derived from K-means Cluster 1, and the intersections account for (76.7% 97.5%) and (22.3% and 70.4%), respectively. K-means Cluster 2 and SOM Cluster 1 are highly overlapping, and the intersection accounts for the ratios of 66.9% and 95.0%. K-means Cluster 3 and SOM Cluster 2 are highly overlaid, and the intersection ratio is 94.1% and 82.1%. For the clustering results of K-means, TB-a and TC-a of Cluster 3 are the highest ($P < 0.001$), TB-a of Cluster 2 is the lowest ($P < 0.001$), and TB-a of Cluster 1 is between Cluster 2 and Cluster 3 ($P < 0.001$). Cluster 1 has the highest TB-b and TC-b ($P < 0.001$), Cluster 2 has the lowest TB-b and TC-b ($P < 0.001$), and TB-b and TC-b of Cluster 3 are between Cluster 1 and Cluster 2 ($P < 0.001$). Cluster 1 has the highest TB-ASM and TC-ASM ($P < 0.001$), Cluster 3 has the lowest TB-ASM and TC-ASM ($P < 0.001$), and TB-ASM and TC-ASM of Cluster 2 are between the Cluster 1 and Cluster 3 ($P < 0.001$). CON, ENT, and MEAN show the opposite trend. Cluster 2 had the highest Per-all ($P < 0.001$). SOM divides K-means Cluster 1 into two categories. There is almost no difference in texture features between Cluster 3 and Cluster 4 in the SOM clustering results. Cluster 3's TB-L, TC-L, and Per-all are lower than Cluster 4 ($P < 0.001$), Cluster 3's TB-a, TC-a, TB-b, TC-b, and Per-part are higher than Cluster 4 ($P < 0.001$). *Conclusions*. The precise tongue image features calculated by TDAS are the basis for characterizing the disease state of diabetic people. Unsupervised learning technology combined with statistical analysis is an important means to discover subtle changes in the tongue features of diabetic people. The machine vision analysis method based on unsupervised machine learning technology realizes the classification of the diabetic population based on fine tongue features. It provides a diagnostic basis for the designated diabetes TCM treatment plan.

1. Introduction

The number of diabetic patients worldwide is increasing rapidly, posing a major threat to human health [1]. Limited by diagnostic methods, current treatment guidelines are difficult to provide individualized treatment plans. Therefore, existing treatment methods are still difficult to prevent the progression of diabetes and the occurrence and

development of complications. Further refined classification of diabetes is conducive to individualized treatment [2]. Syndrome differentiation and treatment of TCM is the practice of individualized treatment. The patient's disease information is collected through tongue diagnosis, pulse diagnosis, and questioning, and the disease is further classified through analysis to formulate an individualized treatment plan [3, 4]. Among them, tongue diagnosis is convenient and noninvasive and can reflect a lot of physiological and pathological information about the human body, so it is an indispensable diagnostic method of TCM [5].

Tongue diagnosis is a classic diagnosis method of TCM. TCM doctors evaluate the disease state by observing changes in tongue body (TB) and tongue coating (TC). They have a long history and rich practical experience. However, on the one hand, due to the influence of light, temperature, humidity, and viewing angle, subtle changes in the color and texture of the tongue require well-trained and experienced physicians to observe. On the other hand, the written tongue diagnosis information is inevitably subjective through visual observation. In order to further improve the standardization and information level of TCM diagnosis, the tongue diagnosis instrument was invented. The tongue diagnosis instrument is equipped with a high-definition camera and a standard light source, which can capture a flat and clear tongue image. It can effectively extract tongue image features with computer image processing technology. These digital tongue features can more accurately characterize the different states of the disease.

Machine learning has been widely used in the field of tongue diagnosis, but the current research focuses on the field of supervised learning, which requires manual calibration of the tongue image [6–8]. It is not difficult to label data with clear diagnostic criteria. However, applying high-dimensional digital tongue features to the classification of the diabetic population is a new research field, and it is difficult to give clear diagnostic criteria, which makes it difficult to explore its internal laws. Unsupervised learning can discover the inherent laws of unlabeled data [9, 10]. Given this, we introduce unsupervised learning technology to study the changing laws of tongue image in different states of the diabetic population and use various methods to evaluate the reliability of the clustering results (Figure 1). Our research will provide an effective analysis method for the classification of the tongue image of diabetes and provide a basis for developing individualized treatment plans.

2. Method

2.1. Study Population. The subjects participating in our study came from Shuguang Hospital affiliated to Shanghai University of TCM and several Shanghai community hospitals. The data collection period is from August 6, 2018, to December 31, 2019. According to the diagnostic criteria issued by ADA 2020 [11], it is determined that subjects with diabetes have fasting blood glucose ≥ 7.0 mmol/L, blood glucose two hours after a meal ≥ 11.1 mmol/L, or HbA1c $\geq 6.5\%$. A total of 598 diabetic subjects were included in the study (Table 1). All subjects signed an informed consent form. The study was approved by the IRB of Shuguang Hospital affiliated with Shanghai University of TCM.

2.2. Tongue Image Collection and Analysis. The intelligent diagnosis technology research team independently developed the TFDA-1 tongue diagnosis instrument (Figure 2) at the Shanghai University of TCM. It is equipped with a standard light source and a high-definition camera, which can take images of the subject's tongue in a stable light environment. The Tongue Diagnosis Analysis System (TDAS) is used to extract features from tongue images. TDAS can realize the segmentation of tongue body (TB) and tongue coating (TC) [12] and calculate the color feature, texture feature, and tongue coating ratio feature of TB and TC, respectively. Color features include RGB, Lab, YCrCb, and other color space features. In order to facilitate data analysis and interpretation, we chose the color feature of the Lab color space. L is used to indicate brightness, the larger the L value, the brighter, and vice versa, the darker; a indicates the range from red to green, a positive value indicates red, a higher value indicates red, a negative value indicates green, and a lower value indicates green; b indicates the range from yellow to blue, with positive values representing yellow, higher values being more yellow, negative values representing blue, and lower values being blue. Texture features include angular second moment (ASM), entropy (ENT), contrast (CON), and mean (MEAN). ASM has the opposite meaning of ENT, CON, and MEAN. The larger the ASM, the smaller the ENT, CON, and MEAN, and the finer the texture of the tongue, and vice versa. Per-all is inversely proportional to Per-part. The larger the Per-all, the smaller the Per-part, the larger the tongue coating area, and vice versa.

2.3. Cluster Analysis. The data analyzed by the clustering algorithm are uncalibrated data. In order to evaluate the reliability of the clustering results, we chose K-means and Self-Organizing Maps (SOM) Network to confirm each other [13, 14]. t-SNE and K-means are implemented and performed in Scikit-learn version 0.23.2 [15].

2.3.1. K-Means Algorithm. Let X be the tongue feature and the dimension be n; then $X = \{x_i : i = 1, \cdots, n\}$. Let the clustering result be C and be divided into k clusters; then $C = \{C_j : j = 1, \cdots, k\}$ [16].

First, k samples $\{\mu_1, \mu_2, \cdots, \mu_k\}$ are randomly selected from X as initial cluster centers. Calculate the distance $d_{ij} = \|x_i - \mu_j\|_2$ from the sample x_i to each cluster center μ_j $(1 < j < k)$. The cluster center closest to the sample is recorded as the category of the sample $\lambda_i = \operatorname{argmin}_{j \in \{1,2,\cdots,k\}} d_{ij}$, and the sample x_i is placed in the corresponding cluster $C_{\lambda_i} = C_{\lambda_i} \cup \{x_i\}$.

Then calculate the new cluster center μ_j', $\mu_j' = (1/C_j)\sum_{x \in C_j} x$. If $\mu_j' \neq \mu_j$, then update μ_j to μ_j'.

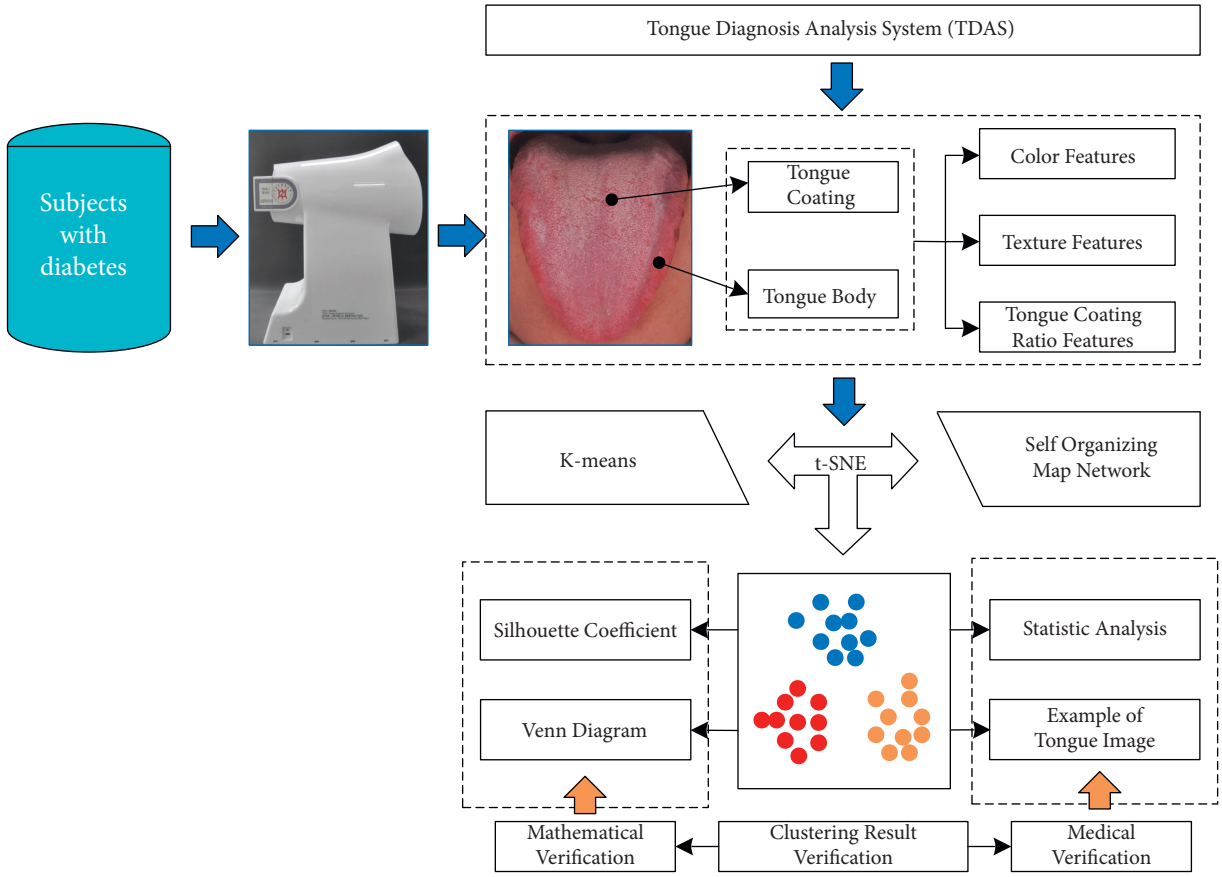

Figure 1: Work flow.

Table 1: Dataset description.

Class	Item	Mean ± SD/median (P_{25}, P_{75})
Basic information	Male (%)	357 (59%)
	Age (years)	58.00 (49.00, 63.00)
Glucose and lipid metabolism	GLU (mmol/L)	7.22 (6.05, 8.50)
	HbA1c (%)	6.75 (6.50, 7.40)
	TCHO (mmol/L)	5.20 (4.54, 5.92)
	TG (mmol/L)	1.79 (1.25, 2.61)
	LDL (mmol/L)	2.91 ± 0.96
	HDL (mmol/L)	1.20 (1.03, 1.41)
Liver function	ALT (U/L)	23 (16, 38)
	AST (U/L)	21 (17, 29)
	GGT (U/L)	32 (21, 51)
Kidney function	UA (umol/L)	338 (287, 411)
	CREA (umol/L)	69 (58, 80)
	BUN (mmol/L)	5.19 (4.47, 6.08)

Note. Laboratory results included the following: (1) glucose and lipid metabolism including blood glucose (GLU), glycosylated hemoglobin (HbA1c), total cholesterol (TCHO), triglycerides (TG), low-density lipoprotein (LDL), and high-density lipoprotein (HDL); (2) liver function including alanine aminotransferase (ALT), aspartate transaminase (AST), and gamma-glutamyl transpeptidase (GGT); (3) kidney function including uric acid (UA), creatinine (CREA), and blood urea nitrogen (BUN).

The *K*-means algorithm needs to specify the number of clusters k value and determine the real k value through the elbow method. As the number of clusters k increases, the division of clusters becomes more and more refined. The sum of the squared errors (SSE) will gradually become smaller (Equation (1)). When k is less than the true number of clusters, the decrease in SSE increases. When k is greater than the number of true clusters, the decrease in SSE tends to

(a) (b) (c)

Figure 2: TFDA-1 tongue diagnosis instrument.

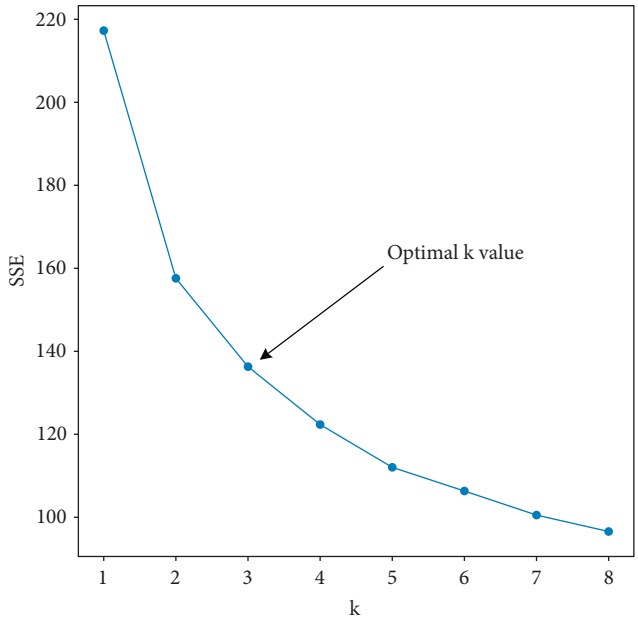

Figure 3: Elbow diagram.

be flat. The relationship graph between SSE and k presents an elbow shape, and the k value corresponding to the elbow is the true number of clusters (Figure 3).

$$\mathrm{SSE} = \sum_{i=1}^{k} \sum_{p \in C_i} \left| p - m_{C_i} \right|^2, \tag{1}$$

where C_i is the i cluster, p is the sample in C_i, and m_{C_i} is the mean value of all samples in C_i.

2.3.2. SOM Networks. SOM can classify uncalibrated data without additional help. Compared with K-means, there is no need to set the number of categories in advance, and the initial impact is small. Introduce the competitive learning mechanism, and determine the winning neuron by calculating the similarity. The spatial position of the output neuron in the topographic map corresponds to the specific domain or feature drawn from the input space. Finally, the input high-dimensional signal is converted into a two-dimensional discrete map [17].

Table 2: Key parameters of t-SNE.

Parameter	Value
Perplexity	30
Early_exaggeration	12
Learning_rate	200
N_iter	1000
Min_grad_norm	$1e{-}7$
Metric	Euclidean
Method	Barnes–Hut

Let the dimension of tongue feature be N and the number of neurons in the computational layer be M. Then the tongue feature is expressed as $X = \{x_i : i = 1, \cdots, N\}$ and the network weight can be expressed as $W_j = \{w_{ji} : j = 1, \cdots, M; i = 1, \cdots, N\}$.

Normalize the input vector and network weight:

$$X' = \frac{X}{\|X\|}, \tag{2}$$

$$W_j' = \frac{W_j}{\left\|W_j\right\|}, \quad 1 < j < M. \tag{3}$$

Input the sample into the network, the sample and the weight vector do a dot product, the output neuron with the largest dot product value wins the competition, and it is recorded as the winning neuron.

$$S_j = W_j \cdot X. \tag{4}$$

Update the neurons in the topological neighborhood of the winning neuron, and renormalize the learned weights.

$$\Delta w_{ji} = \eta(t) \cdot T_{j,I(X)}(t) \cdot \left(x_i - w_{ji}\right). \tag{5}$$

Update the learning rate η and the topological neighborhood so that the distance becomes smaller as time increases. If $\eta \le \eta_{\min}$ or reach the preset number of iterations, the algorithm ends.

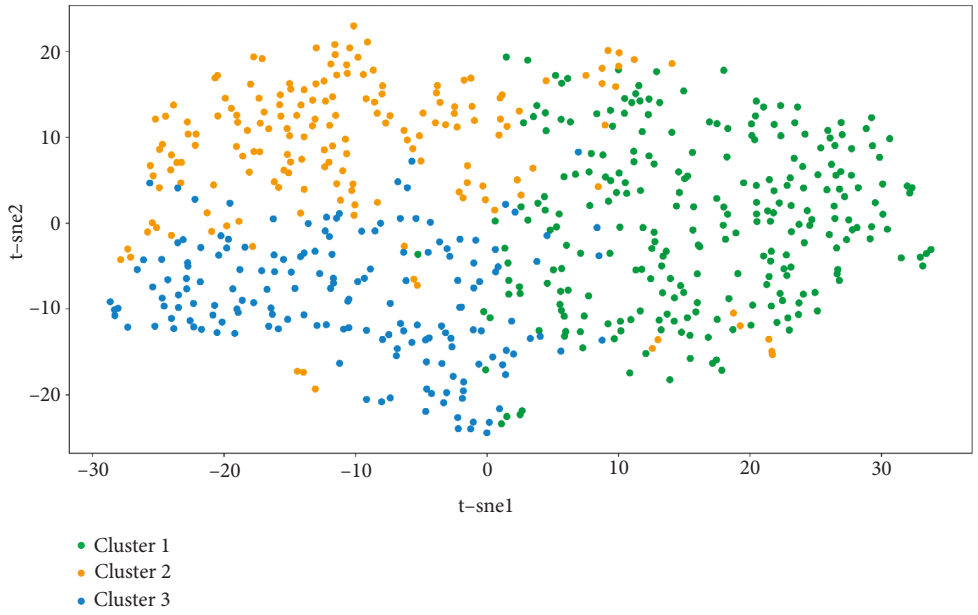

Figure 4: Visualization of *K*-means clustering results based on t-SNE algorithm.

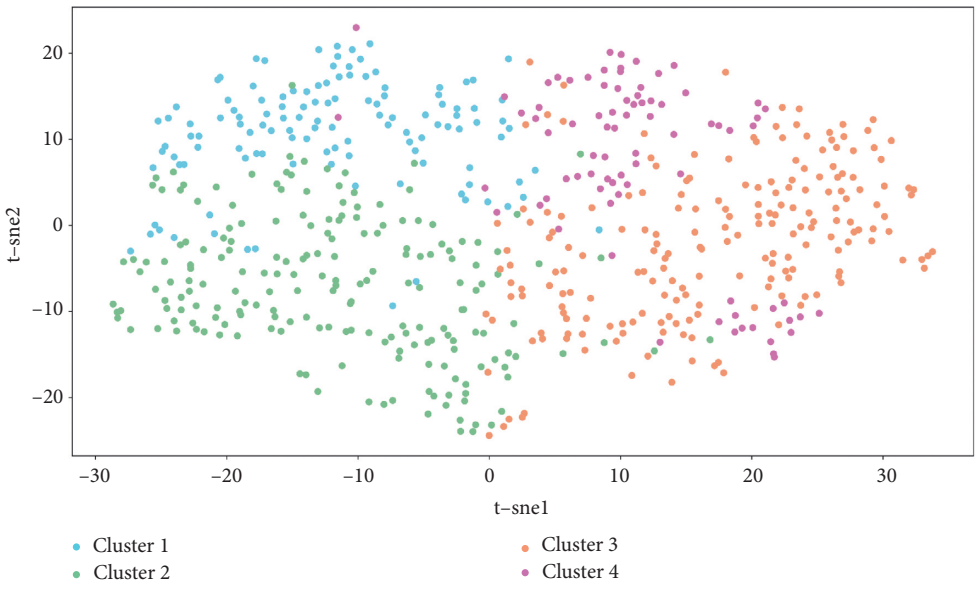

Figure 5: Visualization of SOM clustering results based on t-SNE algorithm.

t-SNE converts the distance relationship into a probability distribution, which can efficiently perform dimensionality reduction calculations for high-dimensional nonlinear tongue feature data. Table 2 shows the key parameters we set for the t-SNE algorithm.

2.4. Statistical Analysis. The statistical analysis program is coded in Python version 3.8.8. Multiple independent samples are compared using the Kruskal–Willis H test in SciPy version 1.4.1 [18], and the post hoc analyses are performed using the Conover test in Scikit-posthocs 0.6.7 [19]. Machine learning and silhouette coefficient are implemented in Scikit-learn version 0.23.2. Statistical image drawing is implemented in Seaborn version 0.11.1 [20] and Matplotlib version 3.2.2 [21].

$$S = \frac{b - a}{\max(a, b)}, \tag{6}$$

where a is the mean distance between a sample and all other points in the same cluster and b is the mean distance between a sample and all other points in the next nearest cluster.

3. Results

The silhouette coefficient of the K-means clustering result is 0.194, and the silhouette coefficient of the SOM clustering

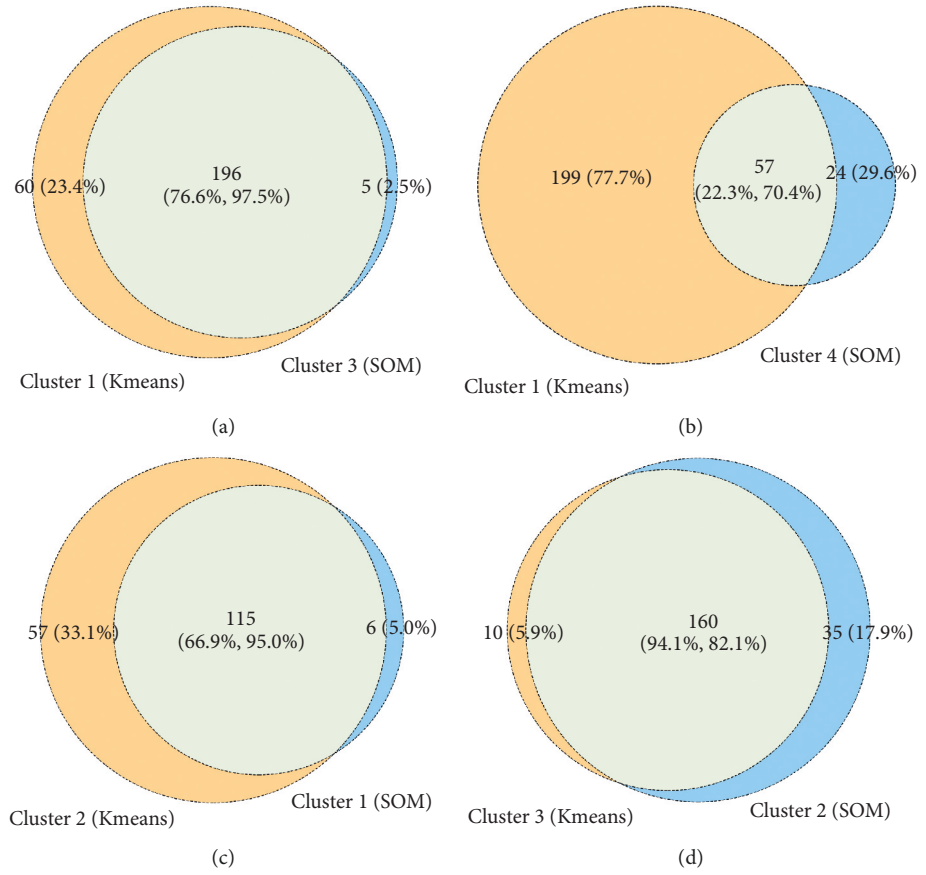

Figure 6: Venn diagram of clustering results of *K*-means and SOM.

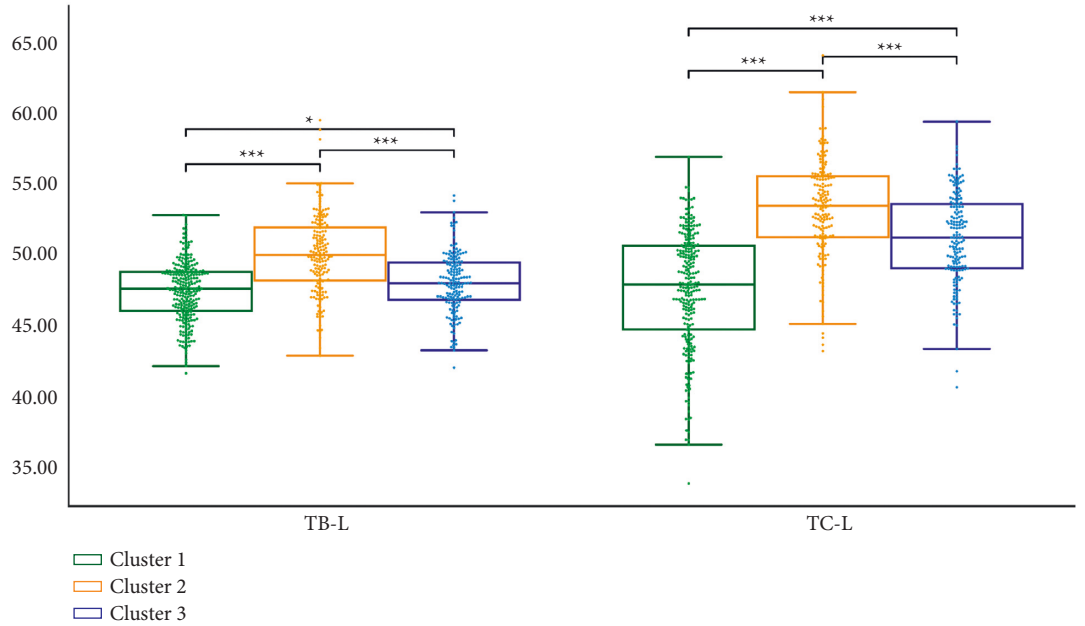

Figure 7: TB-L and TC-L of *K*-means clustering results.

result is 0.127 (Figures 4 and 5). SOM Cluster 3 and Cluster 4 are derived from K-means Cluster 1, and the intersections account for (76.7%, 97.5%) and (22.3%, 70.4%), respectively. K-means Cluster 2 and SOM Cluster 1 are highly overlapped, and the intersection accounts for the ratios are 66.9% and 95.0%. K-means Cluster 3 and SOM Cluster 2 are highly overlapped, and the intersection ratio is 94.1% and 82.1% (Figure 6).

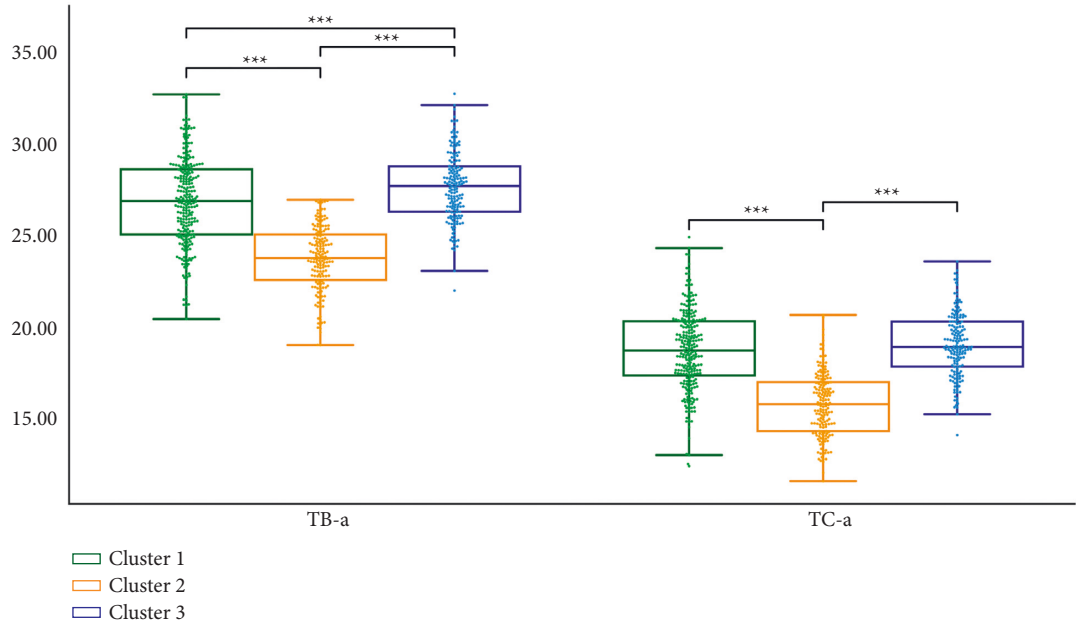

Figure 8: TB-a and TC-a of *K*-means clustering results.

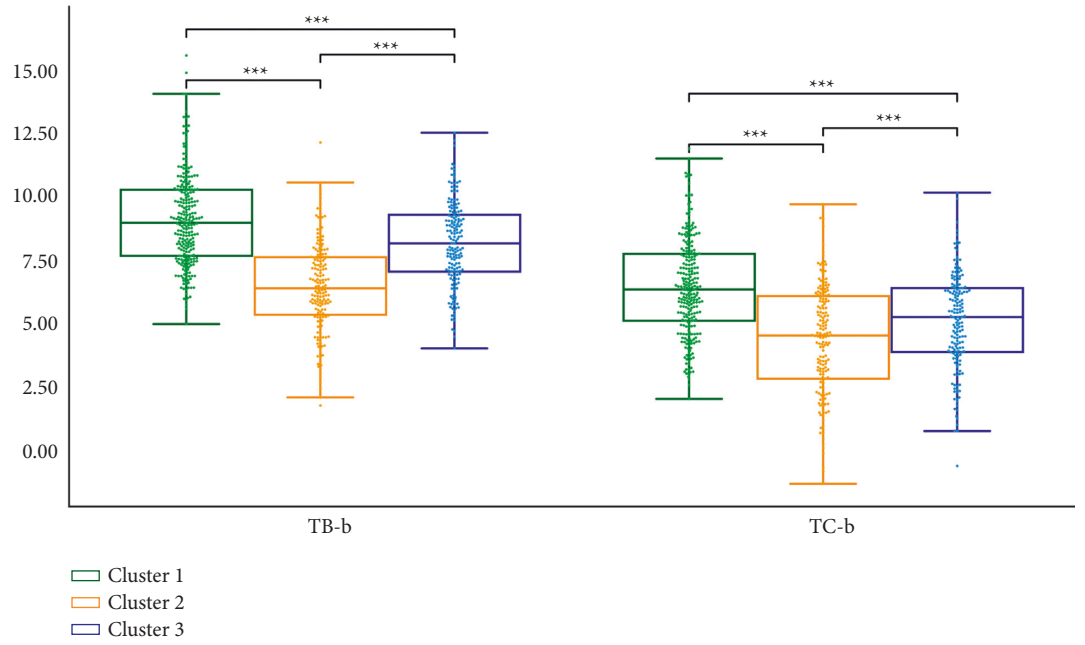

Figure 9: TB-b and TC-b of *K*-means clustering results.

We performed statistical analysis on the clustering results of K-means. Cluster 2 had the highest TB-L and TC-L ($P < 0.001$). Cluster 1 had the lowest TB-L ($P < 0.05$, $P < 0.001$), and Cluster 1 had the lowest TC-L ($P < 0.001$). TB-L of Cluster 3 is between Cluster 1 and Cluster 2 ($P < 0.05$; $P < 0.001$), TB-L of Cluster 3 was between Cluster 1 and Cluster 2 ($P < 0.05$; $P < 0.001$), and the TC-L of Cluster 3 was between Cluster 1 and Cluster 2 ($P < 0.001$) (Figure 7). Cluster 3 had the highest TB-a and TC-a ($P < 0.001$), Cluster 2 had the lowest TB-a ($P < 0.001$), and TB-a of Cluster 1 was between Cluster 2 and Cluster 3 ($P < 0.001$) (Figure 8). Cluster 1 had the highest TB-b and TC-b ($P < 0.001$), Cluster 2 had the lowest TB-b and TC-b ($P < 0.001$), and TB-b and TC-b of Cluster 3 were between Cluster 1 and Cluster 2 ($P < 0.001$) (Figure 9). Cluster 1 had the highest TB-ASM and TC-ASM ($P < 0.001$), Cluster 3 had the lowest TB-ASM and TC-ASM ($P < 0.001$), and TB-ASM and TC-ASM of Cluster 2 were between Cluster 1 and Cluster 3 ($P < 0.001$) (Figure 10). CON, ENT, and MEAN showed the opposite trend (Figures 11–13). Cluster 2 had the highest Per-all ($P < 0.001$). The Per-part of Cluster 2 was the lowest ($P < 0.001$), the Per-part of Cluster 3 was the highest

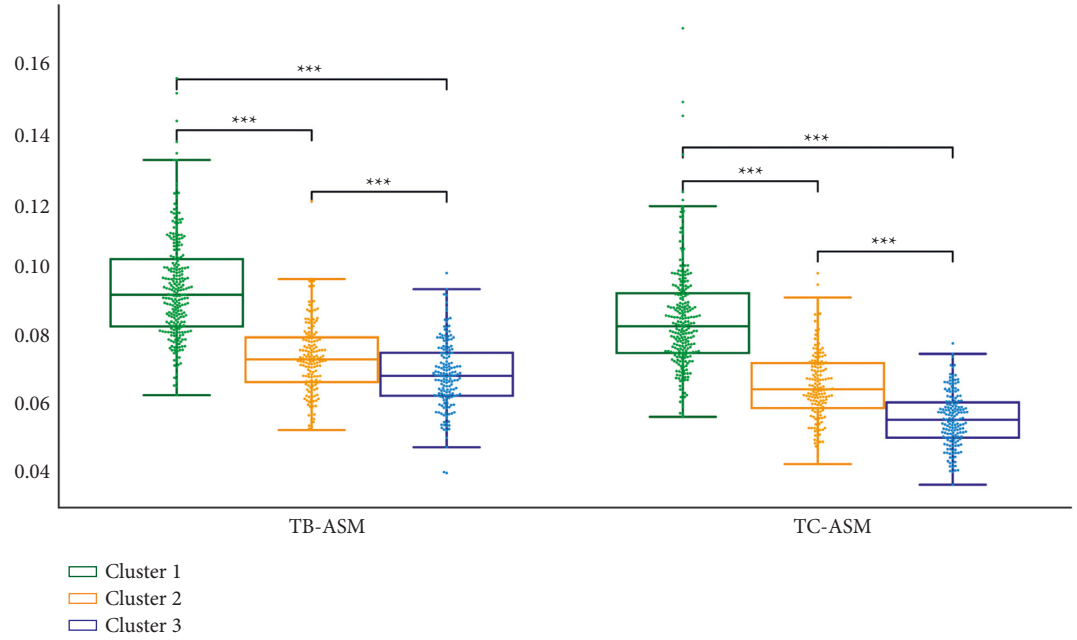

Figure 10: TB-ASM and TC-ASM of *K*-means clustering results.

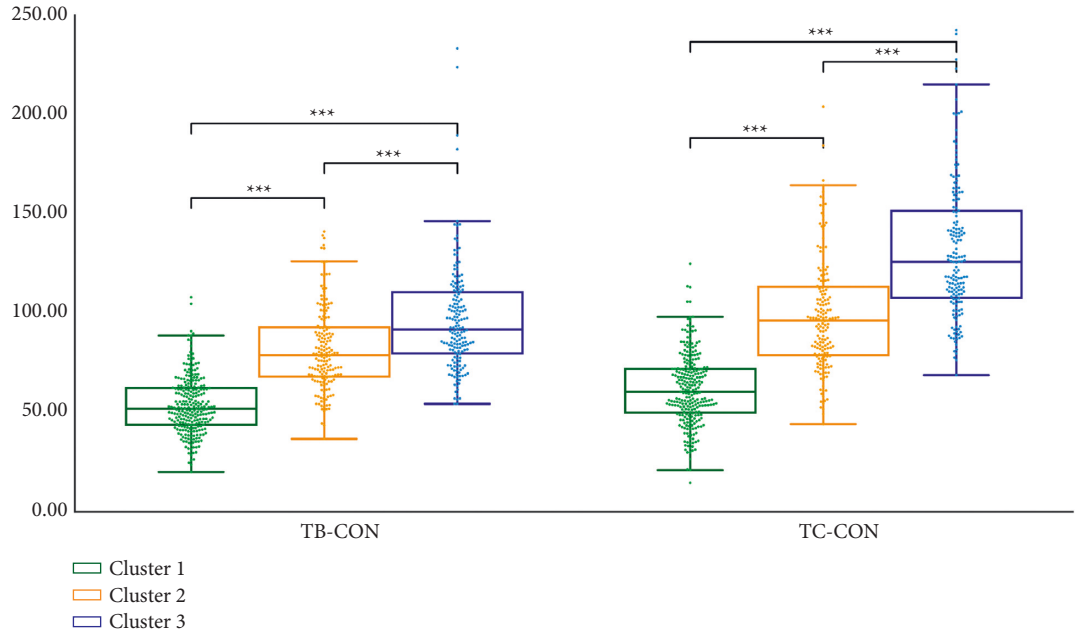

Figure 11: TB-CON and TC-CON of *K*-means clustering results.

($P < 0.01$, $P < 0.001$), and the Per-part of Cluster 1 was between Cluster 2 and Cluster 3 ($P < 0.01$; $P < 0.001$) (Figure 14). Figure 15 shows the representative tongue images in the three groups.

SOM divides K-means Cluster 1 into two categories. Through statistical analysis, it was found that there was almost no difference in texture features between Cluster 3 and Cluster 4 of SOM. Cluster 3's TB-L, TC-L, and Per-all were lower than Cluster 4 ($P < 0.001$), Cluster 3's TB-a, TC-a, TB-b, TC-b, and Per-part were higher than Cluster 4 ($P < 0.001$) (Figure 16 and Table 3). Table 4 shows the results of the statistical analysis of the laboratory tests based on K-means clustering. Table 5 shows the results of the statistical analysis of the laboratory tests based on SOM clustering.

4. Discussion

In order to comprehensively and objectively characterize the tongue image, the tongue features extracted by the TDAS include high-dimensional structural data such as color features, texture features, and tongue coating ratio features.

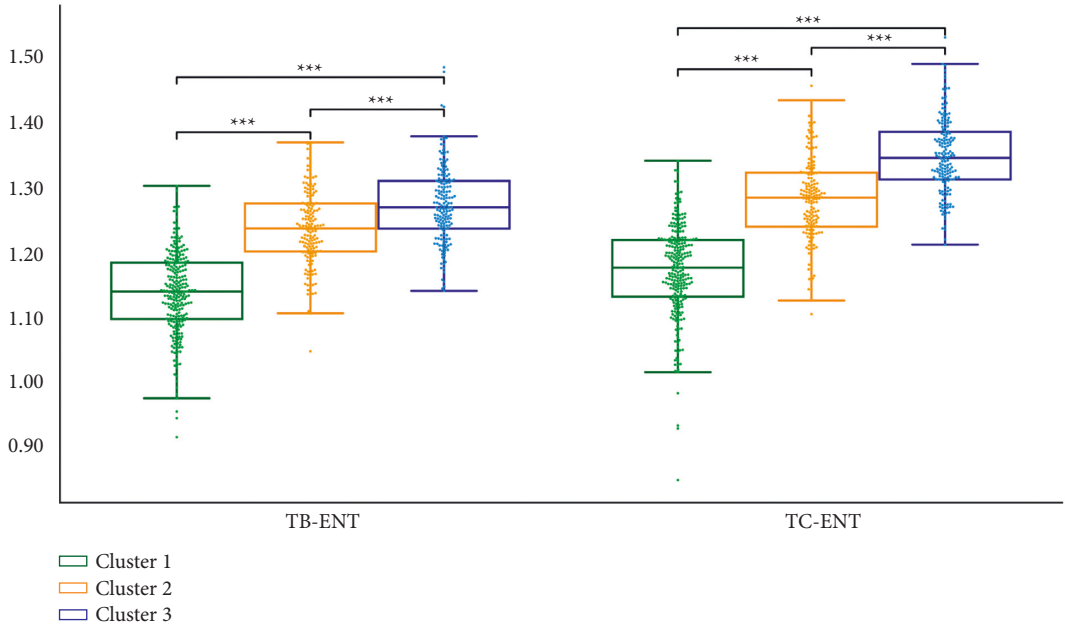

Figure 12: TB-ENT and TC-ENT of *K*-means clustering results.

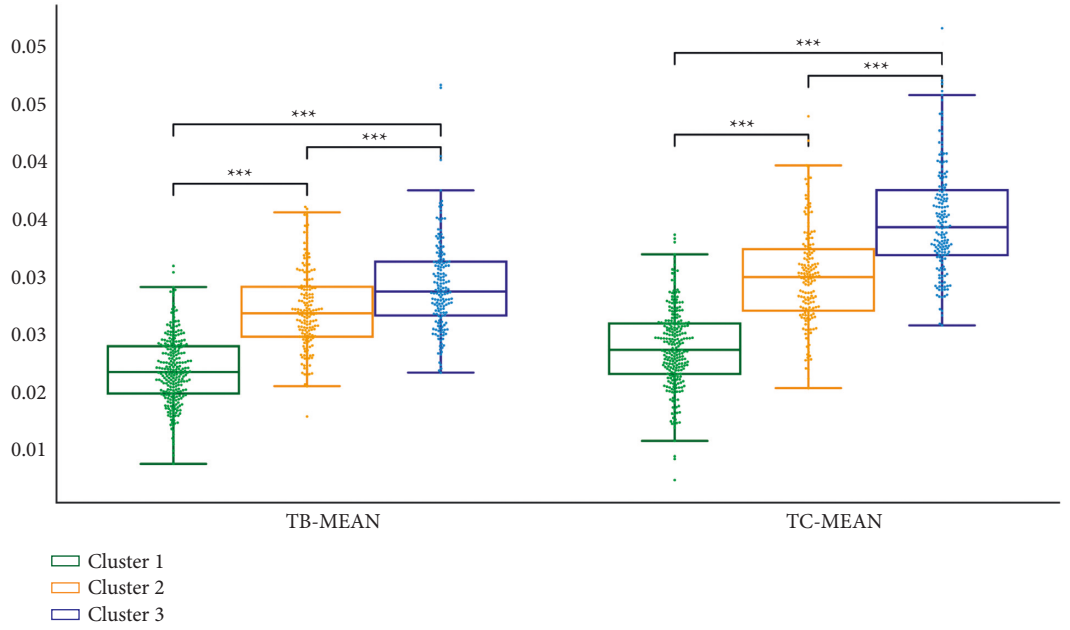

Figure 13: TB-mean and TC-mean of *K*-means clustering results.

These fine tongue features describe tongue changes from multiple angles and explain the tiny differences in tongue features of different states of disease. It is difficult for us to use manual observation and general statistical methods to measure the internal correlation of tongue features and specify diagnostic criteria. Clustering algorithms group samples with similar characteristics into a group through repeated iterations, and are often used to identify outliers, explore the inherent laws of complex data, and make predictions [22, 23].

An adequate evaluation of the clustering results is necessary for reaching a reliable conclusion. We take many measures to evaluate the clustering results. Firstly, K-means and SOM are used to calculate the results of two completely different clustering algorithms and confirm each other. Through the t-SNE algorithm and Venn diagram to visually analyze the clustering results, K-means Clusters and SOM Clusters are highly overlapping. Next, through statistical analysis, it is found that there are extremely significant differences in the tongue features

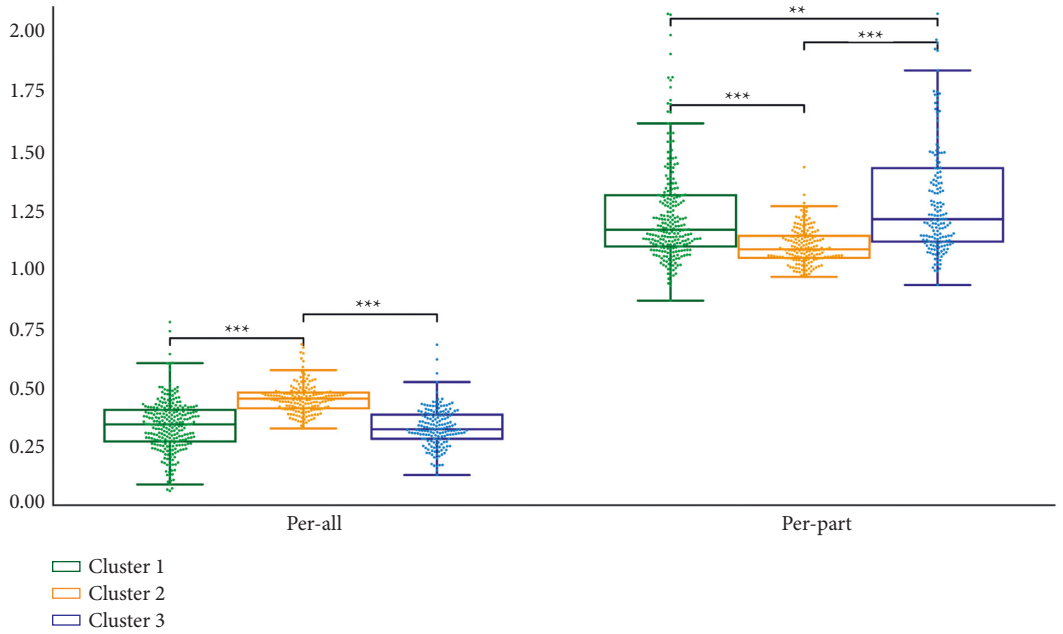

Figure 14: Per-all and per-part of *K*-means clustering results.

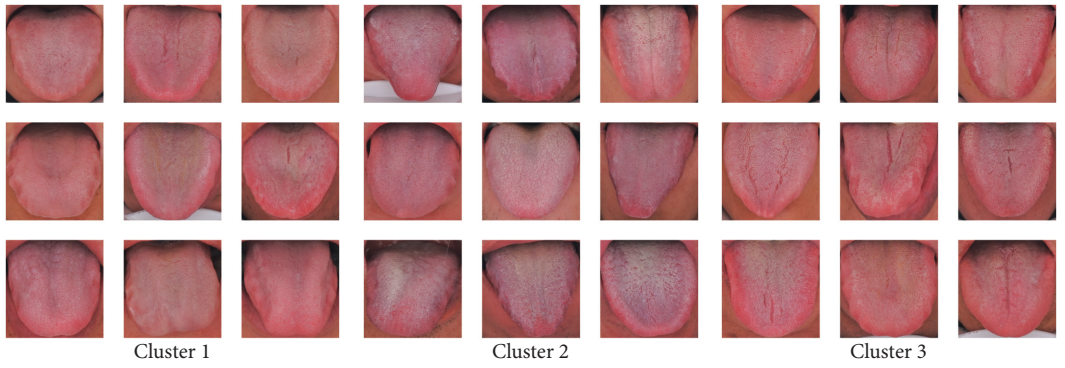

Figure 15: Examples of tongue images of *K*-means' 3 clusters. Note. To the naked eye, there are significant differences between the three types of tongue images. (a) Cluster 1 has the most delicate texture, (b) Cluster 2 has the most purple color, and (c) Cluster 3 has the roughest texture.

between the clusters, which proves that the differences between the clusters are extremely large. Finally, by displaying the representative tongue samples of each cluster, it is found that the differences are visible to the naked eye.

The classification of the diabetic population based on the tongue features provides a new understanding of the heterogeneity of diabetes. We found that the tongues of people with diabetes are divided into three types through cluster analysis. The first type of tongue is characterized by a red tongue and a dry and rough tongue; the second type of tongue is characterized by a purple tongue and a large and thick tongue coating; the third type of tongue is mainly characterized by its fine texture. Chinese medicine believes that the first tongue is a sign of heat syndrome, the second tongue is a sign of phlegm and blood stasis, and the third tongue is a sign of deficiency syndrome. The above classification is expected to provide an objective basis for the individualized treatment of diabetes and has potential clinical value. K-means divides the diabetic population into 3 categories, while SOM divides the diabetic population into 4 categories, of which two clusters are highly overlapping. SOM subdivides one cluster in K-means into two categories. The difference between the two categories is mainly reflected in the color features and coating ratio features. Our research proves that digital tongue features can characterize subtle changes in the tongue of diabetic people. High-precision image measurement technology combined with unsupervised machine learning technology is the key means to discover this difference.

Our research found that machine vision can keenly perceive the subtle changes in the tongue of diabetic patients and realize the fine classification of the diabetic population, providing a diagnosis basis for the individualized treatment of diabetes in TCM. We verify each other through two clustering algorithms with completely different mechanisms.

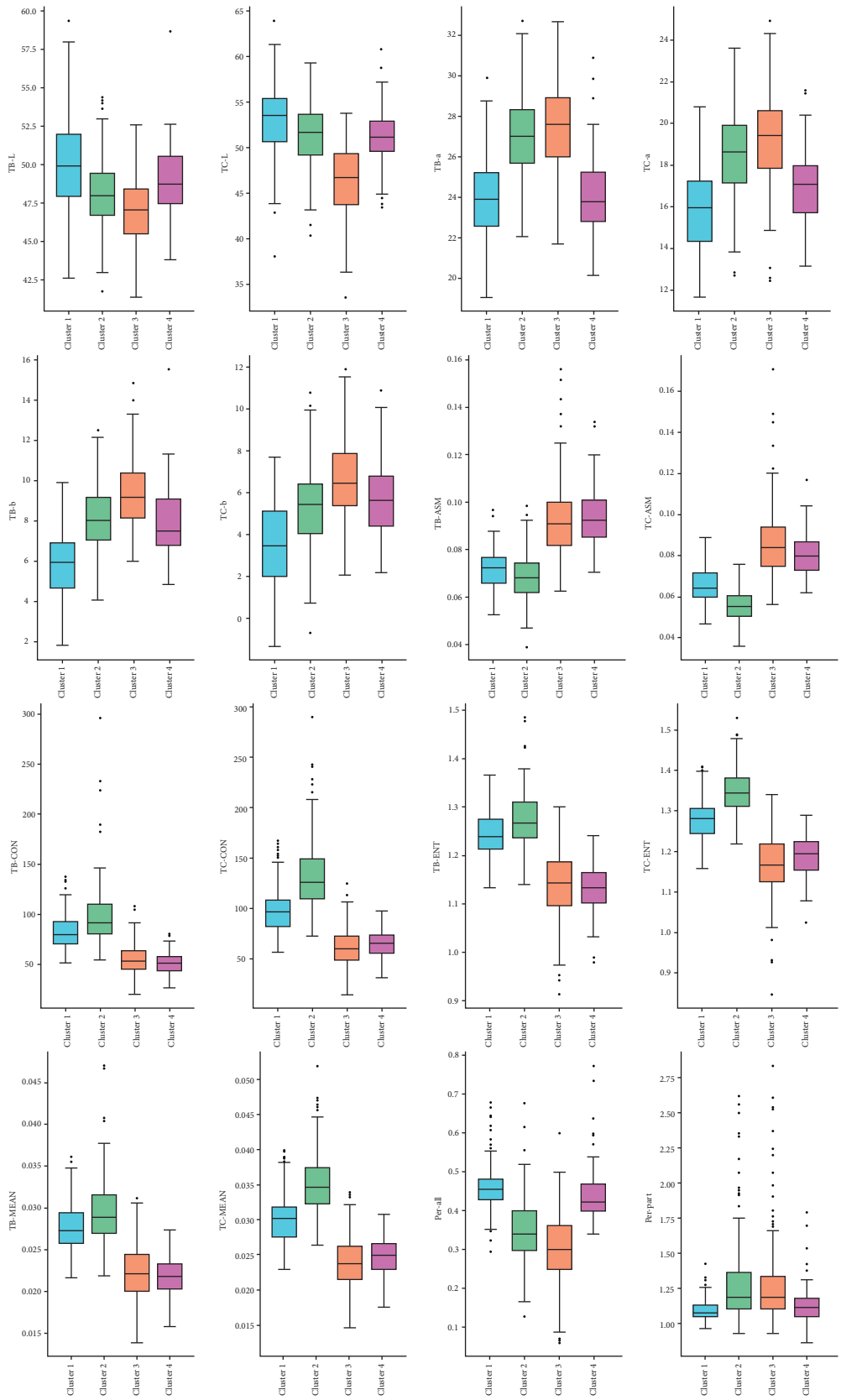

Figure 16: Tongue feature distribution based on SOM algorithm.

Table 3: Tongue feature comparation based on SOM 4 clusters.

Item	Cluster 1	Cluster 2	Cluster 3	Cluster 4
TB-L	49.905 (47.923, 51.967)	47.959 (46.683, 49.431)***	47.057 (45.501, 48.402)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	48.706 (47.468, 50.538)$^{\triangle\square\square\star\star\star}$
TB-a	23.918 (22.537, 25.220)	27.014 (25.663, 28.404)***	27.660 (25.975, 28.907)$^{\#\#\#\blacktriangle}$	23.780 (22.818, 25.239)$^{\square\square\square\star\star\star}$
TB-b	5.949 (4.614, 6.866)	7.999 (7.065, 9.162)***	9.174 (8.118, 10.371)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	7.478 (6.750, 9.075)$^{\triangle\triangle\triangle\star\star\star}$
TC-L	53.518 (50.644, 55.435)	51.717 (49.192, 53.752)***	46.691 (43.692, 49.318)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	51.142 (49.669, 52.920)$^{\triangle\triangle\triangle\star\star\star}$
TC-a	15.905 ± 1.915	$18.518 \pm 2.043^{***}$	$19.220 \pm 2.189^{\#\#\#\blacktriangle\blacktriangle}$	$16.948 \pm 1.833^{\triangle\triangle\triangle\square\square\square\star\star\star}$
TC-b	3.575 ± 2.035	$5.304 \pm 1.863^{***}$	$6.590 \pm 1.854^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	$5.730 \pm 1.811^{\triangle\triangle\triangle\star\star\star}$
TB-CON	80.025 (70.801, 92.922)	91.638 (80.664, 110.948)***	53.213 (45.052, 63.684)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	51.768 (44.479, 58.037)$^{\triangle\triangle\triangle\square\square\square\star}$
TB-ASM	0.072 (0.066, 0.077)	0.068 (0.062, 0.074)**	0.091 (0.082, 0.100)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	0.092 (0.085, 0.101)$^{\triangle\triangle\triangle\square\square\square}$
TB-ENT	1.238 (1.212, 1.275)	1.268 (1.236, 1.310)***	1.143 (1.096, 1.188)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	1.134 (1.101, 1.164)$^{\triangle\triangle\triangle\square\square\square}$
TB-MEAN	0.027 (0.026, 0.029)	0.029 (0.027, 0.032)***	0.022 (0.020, 0.024)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	0.022 (0.020, 0.023)$^{\triangle\triangle\triangle\square\square\square}$
TC-CON	96.448 (82.107, 107.996)	125.628 (109.574, 149.131)***	60.078 (48.983, 72.450)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	65.916 (55.146, 74.215)$^{\triangle\triangle\triangle\square\square\square}$
TC-ASM	0.064 (0.060, 0.072)	0.055 (0.050, 0.060)***	0.084 (0.075, 0.094)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	0.080 (0.073, 0.087)$^{\triangle\triangle\triangle\square\square\square}$
TC-ENT	1.282 (1.245, 1.306)	1.343 (1.312, 1.381)***	1.166 (1.125, 1.218)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	1.194 (1.152, 1.225)$^{\triangle\triangle\triangle\square\square\square}$
TC-MEAN	0.030 (0.028, 0.032)	0.035 (0.032, 0.037)***	0.024 (0.021, 0.026)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	0.025 (0.023, 0.027)$^{\triangle\triangle\triangle\square\square\square}$
Per-all	0.456 (0.429, 0.481)	0.341 (0.297, 0.400)***	0.303 (0.250, 0.363)$^{\#\#\#\blacktriangle\blacktriangle\blacktriangle}$	0.422 (0.400, 0.468)$^{\triangle\square\square\square\star\star\star}$
Per-part	1.076 (1.042, 1.132)	1.188 (1.099, 1.363)***	1.190 (1.106, 1.338)$^{\#\#\#}$	1.112 (1.044, 1.177)$^{\square\square\square\star\star\star}$

*denotes the comparison of Cluster 1 and Cluster 2, $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.$^{\#}$ denotes the comparison of Cluster 1 and Cluster 3, $^{\#}P < 0.05$,$^{\#\#}P < 0.01$, and $^{\#\#\#}P < 0.001$. $^{\triangle}$ denotes the comparison of Cluster 1 and Cluster 4, $^{\triangle}P < 0.05$, $^{\triangle\triangle}P < 0.01$, and $^{\triangle\triangle\triangle}P < 0.001$. $^{\blacktriangle}$ denotes the comparison of Cluster 2 and Cluster 3, $^{\blacktriangle}P < 0.05$, $^{\blacktriangle\blacktriangle}P < 0.01$, and $^{\blacktriangle\blacktriangle\blacktriangle}P < 0.001$. $^{\square}$ denotes the comparison of Cluster 2 and Cluster 4, $^{\square}P < 0.05$, $^{\square\square}P < 0.01$, and $^{\square\square\square}P < 0.001$. * denotes the comparison of Cluster 3 and Cluster 4, $^{\star}P < 0.05$, $^{\star\star}P < 0.01$, and $^{\star\star\star}P < 0.001$.

Table 4: Laboratory results (K-means).

Item	Cluster 1 ($n = 256$)	Cluster 2 ($n = 172$)	Cluster 3 ($n = 170$)
Gender (n, %)	144 (56)	112 (65)	86 (51)
Age (years)	57 (49, 63)	57 (49, 63)	60 (52, 65)$^{\star\bigstar}$
GLU (mmol/L)	7.300 (6.215, 8.412) ($n = 256$)	7.480 (6.395, 9.045) ($n = 171$)	6.845 (5.592, 8.100)$^{\star\bigstar\bigstar}$ ($n = 170$)
TCHO (mmol/L)	5.306 ± 1.106 ($n = 244$)	5.284 ± 1.066 ($n = 163$)	5.118 ± 1.051 ($n = 164$)
TG (mmol/L)	1.730 (1.240, 2.420) ($n = 244$)	1.920 (1.330, 2.960) ($n = 163$)	1.660 (1.222, 2.468) ($n = 164$)
LDL (mmol/L)	2.935 (2.328, 3.652) ($n = 212$)	2.915 (2.260, 3.485) ($n = 134$)	2.890 (2.165, 3.485) ($n = 147$)
HDL (mmol/L)	1.185 (1.050, 1.455) ($n = 212$)	1.190 (1.002, 1.300) ($n = 134$)	1.230 (1.030, 1.400) ($n = 147$)
UA (umol/L)	333.500 (280.750, 412.750) ($n = 256$)	341.000 (288.000, 421.000) ($n = 171$)	346.500 (293.750, 399.750) ($n = 170$)
ALT (U/L)	23.000 (15.000, 40.000) ($n = 256$)	24.000 (16.500, 35.500) ($n = 171$)	24.000 (15.250, 38.000) ($n = 170$)
AST (U/L)	21.000 (17.000, 30.000) ($n = 252$)	21.000 (17.000, 27.000) ($n = 167$)	21.000 (18.000, 29.000) ($n = 169$)
GGT (U/L)	31.000 (21.000, 51.750) ($n = 250$)	35.500 (24.250, 51.750) ($n = 162$)	31.000 (21.000, 46.000) ($n = 161$)
TBIL (umol/L)	12.400 (9.000, 16.200) ($n = 221$)	13.000 (10.350, 16.775) ($n = 146$)	11.400 (9.400, 15.500) ($n = 141$)
CREA (umol/L)	68.500 (57.000, 80.000) ($n = 256$)	71.000 (60.000, 80.000) ($n = 171$)	70.000 (59.000, 80.000) ($n = 170$)
BUN (mmol/L)	5.055 (4.400, 6.000) ($n = 254$)	5.275 (4.500, 6.250) ($n = 170$)	5.190 (4.600, 6.000) ($n = 170$)
eGFR (ml/min∗1.73 m̂2)	107.900 (92.200, 121.100) ($n = 217$)	104.220 (92.240, 117.500) ($n = 137$)	101.000 (86.000, 114.900) ($n = 138$)

*denotes the comparison of Cluster 1 and Cluster 3, $^{\star}P < 0.05$, $^{\star\star}P < 0.01$, and $^{\star\star\star}P < 0.001$. $^{\bigstar}$ denotes the comparison of Cluster 2 and Cluster 3, $^{\bigstar}P < 0.05$, $^{\bigstar\bigstar}P < 0.01$, and $^{\bigstar\bigstar\bigstar}P < 0.001$.

Table 5: Laboratory results (SOM).

Item	Cluster 1 ($n = 121$)	Cluster 2 ($n = 195$)	Cluster 3 ($n = 201$)	Cluster 4 ($n = 81$)
Gender (n, %)	82 (68)	99 (51)	112 (56)	49 (60)
Age (years)	56 (48, 62)	61 (54, 65)**	57 (50, 63)▲	54 (46, 61)□□□
GLU (mmol/L)	7.565 (6.360, 9.250) ($n = 120$)	6.900 (5.585, 8.100)** ($n = 195$)	7.200 (6.330, 8.240) ($n = 201$)	7.810 (6.300, 9.670)□□ ($n = 81$)
TCHO (mmol/L)	5.376 ± 1.128 ($n = 114$)	5.061 ± 1.022 ($n = 189$)	5.296 ± 1.084 ($n = 192$)	5.384 ± 1.098 ($n = 76$)
TG (mmol/L)	1.935 (1.397, 2.860) ($n = 114$)	1.610 (1.200, 2.420) ($n = 189$)	1.745 (1.240, 2.433) ($n = 192$)	1.975 (1.488, 2.832) ($n = 76$)
LDL (mmol/L)	2.960 (2.230, 3.580) ($n = 95$)	2.790 (2.210, 3.370) ($n = 167$)	2.940 (2.335, 3.570) ($n = 164$)	2.990 (2.310, 3.705) ($n = 67$)
HDL (mmol/L)	1.190 (1.025, 1.365) ($n = 95$)	1.230 (1.020, 1.440) ($n = 167$)	1.180 (1.048, 1.450) ($n = 164$)	1.110 (1.025, 1.325) ($n = 67$)
UA (umol/L)	345.(294, 419) ($n = 120$)	337 (292, 401) ($n = 195$)	340 (287, 412) ($n = 201$)	326 (253, 422) ($n = 81$)
ALT (U/L)	25 (17, 38) ($n = 120$)	23 (16, 36) ($n = 195$)	23 (16, 40) ($n = 201$)	24 (15, 35) ($n = 81$)
AST (U/L)	21 (18, 29) ($n = 119$)	21 (18, 28) ($n = 191$)	21 (17, 30) ($n = 198$)	20 (17, 29) ($n = 80$)
GGT (U/L)	36 (24, 53) ($n = 115$)	31 (21, 44) ($n = 182$)	31 (21, 53) ($n = 195$)	37 (24, 57) ($n = 81$)
TBIL (umol/L)	13 (11, 17) ($n = 104$)	12 (9, 16) ($n = 161$)	12 (9, 16) ($n = 173$)	13 (9, 17) ($n = 70$)
CREA (umol/L)	72 (61, 80) ($n = 120$)	70 (59, 81) ($n = 195$)	68 (56, 79) ($n = 201$)	71 (57, 80) ($n = 81$)
BUN (mmol/L)	5.295 (4.470, 6.455) ($n = 120$)	5.120 (4.518, 5.975) ($n = 194$)	5.100 (4.415, 6.000) ($n = 199$)	5.000 (4.450, 6.100) ($n = 81$)
eGFR (ml/min∗1.73 m̂2)	104 (93, 117) ($n = 99$)	102 (87, 117) ($n = 155$)	107 (91, 122) ($n = 170$)	103 (92, 115) ($n = 68$)

∗ denotes the comparison of Cluster 1 and Cluster 2, $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. ▲ denotes the comparison of Cluster 2 and Cluster 3, $^{▲}P < 0.05$, $^{▲▲}P < 0.01$, and $^{▲▲▲}P < 0.001$. □ denotes the comparison of Cluster 2 and Cluster 4, $^{□}P < 0.05$, $^{□□}P < 0.01$, and $^{□□□}P < 0.001$.

Statistical analysis, t-SNE, Venn diagram visualization analysis, and tongue image examples fully prove the reliability of the clustering results. However, our data lack information on the course of the disease, and it is impossible to further infer the evolutionary law of tongue features in the course of diabetes.

5. Conclusion

The fine tongue features calculated by TDAS are the basis for classifying the diabetic population. We have discovered the inherent law of tongue features of people with diabetes through unsupervised machine learning technology. The application of statistical analysis to compare the differences between the clusters clarified the meaning of the tongue features of each cluster. Our study laid the foundation for implementing individualized treatment of diabetes in TCM. In the future, we will expand the sample size and build a standard tongue feature database for diabetic population classification.

Data Availability

The experimental data used to support the findings of this study were supplied by Prof. Jiatuo Xu under license and thus cannot be made freely available. Requests for access to these data should be made to Prof. Jiatuo Xu (e-mail: xjt@fudan.edu.cn).

Disclosure

Jun Li and Longtao Cui are joint first authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Jun Li drafted the article and performed data analysis. Sihan Wang, Yulin Shi, and Jiayi Liu contributed to data collection and data analysis. Xiaojuan Hu, Liping Tu, Jingbin Huang, and Longtao Cui contributed to the revision of the article. Changle Zhou, Yongzhi Li, Longtao Cui, Jingbin Huang, and Jiatuo Xu contributed to the conception of the work. All authors contributed to the interpretation of data and critical revision of the article. All authors gave final approval of the version to be published.

Acknowledgments

The authors thank every subject who participated in the study. The authors gratefully acknowledge the financial supports by the National Key Research and Development Program of China (grant numbers: 2017YFC1703300 and 2017YFC1703301), National Natural Science Foundation of China (grant numbers: 81873235, 81973750, 82004258, 81904094, 82104736, and 82104738), Shanghai Science and Technology Commission (grant number: 21010504400), Shanghai Municipal Health Commission (grant number: 2020JQ003), and 1226 Engineering Health Key Project (grant number: BWS17J028).

References

[1] E. Eto, J. Maki, S. Tamada et al., "Assessment of resting energy expenditure and body composition in Japanese pregnant women with diabetes," *Journal of Diabetes Investigation*, vol. 9, no. 4, pp. 959–966, 2018.

[2] E. Ahlqvist, P. Storm, A. Käräjämäki et al., "Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables," *Lancet Diabetes & Endocrinology*, vol. 6, no. 5, pp. 361–369, 2018.

[3] Y. Ma, M. Chen, Y. Guo et al., "Prevention and treatment of infectious diseases by traditional Chinese medicine: a commentary," *Acta Pathologica, Microbiologica et Immunologica Scandinavica*, vol. 127, no. 5, pp. 372–384, 2019.

[4] J. Guo, H. Chen, J. Song, J. Wang, L. Zhao, and X. Tong, "Syndrome differentiation of diabetes by the traditional Chinese medicine according to evidence-based medicine and expert consensus opinion," *Evidence-based Complementary and Alternative Medicine*, vol. 2014, Article ID 492193, 2014.

[5] X. Wang, J. Liu, C. Wu et al., "Artificial intelligence in tongue diagnosis: using deep convolutional neural network for recognizing unhealthy tongue with tooth-mark," *Computational and Structural Biotechnology Journal*, vol. 18, pp. 973–980, 2020.

[6] H. Li, G. Wen, and H. Zeng, "Natural tongue physique identification using hybrid deep learning methods," *Multimedia Tools and Applications*, vol. 78, no. 6, pp. 6847–6868, 2019.

[7] J. Hou, H. Su, B. Yan, H. Zheng, Z. Sun, and X. Cai, "Classification of tongue color based on CNN," in *Proceedings of the IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, Beijing, China, 2017.

[8] X. Lin, Z. Yu, Z. Li, and W. Liu, "Machine learning based tongue image recognition for diabetes diagnosis," in *Proceedings of the International Conference on Machine Learning for Cyber Security*, Guangzhou, China, 2020.

[9] S. L. Nolin, E. Napoli, A. Flores, R. J. Hagerman, and C. Giulivi, "Deficits in prenatal serine biosynthesis underlie the mitochondrial dysfunction associated with the autism-linked FMR1 gene," *International Journal of Molecular Sciences*, vol. 22, no. 11, p. 5886, 2021.

[10] S. Han, E. Taralova, C. Dupre, and R. Yuste, "Comprehensive machine learning analysis of Hydra behavior reveals a stable basal behavioral repertoire," *Elife*, vol. 7, Article ID e32605, 2018.

[11] American Diabetes Association, "Classification and diagnosis of diabetes: standards of medical care in diabetes-2020," *Diabetes Care*, vol. 43, pp. S14–S31, 2020.

[12] J. Xu, L. Tu, Z. Zhang, L. Zhang, and C. Zhou, "The region partition of quality and coating for tongue image based on color image segmentation method," in *Proceedings of the IEEE International Symposium on IT in Medicine and Education*, Piscataway, NJ, USA, 2008.

[13] B. Zheng, S. W. Yoon, and S. S. Lam, "Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1476–1482, 2014.

[14] A. Zadeh Shirazi, S. J. Seyyed Mahdavi Chabok, and Z. Mohammadi, "A novel and reliable computational intelligence system for breast cancer detection," *Medical, & Biological Engineering & Computing*, vol. 56, no. 5, pp. 721–732, 2018.

[15] F. Pedregosa, G. Varoquaux, A. Gramfort et al., "Scikit-learn: machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[16] Z. Zhou, "Cluster," in *Machine Learning*p. 203, Tsinghua University Press, Beijing, China, 2016.

[17] J. A. Bullinaria, "Self organizing maps: fundamentals," *Introduction to Neural Networks*, vol. 16, pp. 1–15, 2004.

[18] P. Virtanen, R. Gommers, T. E. Oliphant et al., "SciPy 1.0: fundamental algorithms for scientific computing in Python," *Nature Methods*, vol. 17, no. 3, pp. 261–272, 2020.

[19] M. A. Terpilowski, "scikit-posthocs: pairwise multiple comparison tests in Python," *Journal of Open Source Software*, vol. 4, no. 36, p. 1169, 2019.

[20] M. L. Waskom, "Seaborn: statistical data visualization," *Journal of Open Source Software*, vol. 6, no. 60, p. 3021, 2021.

[21] N. Ari and M. Ustazhanov, "Matplotlib in python," in *Proceedings of the 11th International Conference on Electronics, Computer and Computation (ICECCO)*, Abuja, Nigeria, 2014.

[22] Y. Liang, A. Han, L. Chai, and H. Zhi, "Using the machine learning method to study the environmental footprints embodied in Chinese diet," *International Journal of Environmental Research and Public Health*, vol. 17, no. 19, p. 7349, 2020.

[23] J. Zhao, Z. Li, Q. Gao et al., "A review of statistical methods for dietary pattern analysis," *Nutrition Journal*, vol. 20, no. 1, p. 37, 2021.