Research Article

Automated Tongue Feature Extraction for ZHENG Classification in Traditional Chinese Medicine

Ratchadaporn Kanawong,1 Tayo Obafemi-Ajayi,1 Tao Ma,2 Dong Xu,1 Shao Li,2 and Ye Duan1

1 Department of Computer Science and Informatics Institute, University of Missouri, Columbia, MO 65211, USA
2 MOE Key Laboratory of Bioinformatics and Bioinformatics Division, TNLST/Department of Automation, Tsinghua University, Beijing 100084, China

Correspondence should be addressed to Shao Li, shaoli@mail.tsinghua.edu.cn and Ye Duan, duanye@missouri.edu

Received 13 February 2012; Accepted 30 March 2012

1. Introduction

Traditional Chinese Medicine (TCM) has a long history in the treatment of various diseases in East Asian countries and is also a complementary and alternative medical system in Western countries. TCM takes a holistic approach to medicine with emphasis on the integrity of the human body and the close relationship between a human and its social and natural environment [1]. TCM applies different therapeutic methods to enhance the body’s resistance to diseases and prevention. TCM diagnosis is based on the information obtained from four diagnostic processes, that is, looking, listening, and smelling, asking, and touching. The most common tasks are taking the pulse and inspecting the tongue [2]. For thousands of years, Chinese medical practitioners have diagnosed the health status of a patients’ internal organs by inspecting the tongue, especially the patterns on the tongue’s surface. The tongue mirrors the viscera. The changes of tongue can objectively manifest the states of a disease, which can help differentiate syndromes, establish treatment methods, prescribe herbs, and determine prognosis of disease.

ZHENG (TCM syndrome) is an integral and essential part of TCM theory. It defines the theoretical abstraction of the symptom profiles of individual patients and thus, used as a guideline in disease classification in Chinese medicine. For example, patients suffering from gastritis may be classified as Cold or Hot ZHENG, whereas patients with different diseases may be classified under the same ZHENG. Tongue appearance is a valuable diagnostic tool for determining ZHENG in patients. In this paper, we explore new modalities for the clinical characterization of ZHENG using various supervised machine learning algorithms. We propose a novel-color-space-based feature set, which can be extracted from tongue images of clinical patients to build an automated ZHENG classification system. Given that Chinese medical practitioners usually observe the tongue color and coating to determine a ZHENG type and to diagnose different stomach disorders including gastritis, we propose using machine-learning techniques to establish the relationship between the tongue image features and ZHENG by learning through examples. The experimental results obtained over a set of 263 gastritis patients, most of whom suffering Cold Zheng or Hot ZHENG, and a control group of 48 healthy volunteers demonstrate an excellent performance of our proposed system.
suffering from the same disease may be grouped into different ZHENGs, whereas different diseases may be grouped as the same ZHENG. The Cold ZHENG (Cold syndrome) and the Hot ZHENG (Hot syndrome) are the two key statuses of ZHENG [3]. Other ZHENGs include Shen-Yang-Xu ZHENG (Kidney-Yang deficiency syndrome), Shen-Xu ZHENG (Kidney deficiency syndrome), and Xue-Yu ZHENG (Blood Stasis syndrome) [4].

In this paper, we explore new modalities for the clinical characterization of ZHENG using various supervised machine-learning algorithms. Using an automated tongue-image diagnosis system, we extract objective features from tongue images of clinical patients and analyze the relationship with their corresponding ZHENG data and disease diagnosis (specifically stomach disorders, i.e., gastritis) obtained from clinical practitioners. We propose a system that learns from the clinical practitioner’s subjective data on how to classify a patient’s health status by extracting meaningful features from tongue images using a rich set of features based on color-space models. Our premise is that Chinese medical practitioners usually observe the tongue color and coating to determine ZHENG such as Hot or Cold ZHENG, and to diagnose different stomach disorders including gastritis. Hence, we propose using machine-learning techniques to establish the relationship between the tongue image features and the ZHENG by learning through examples. We are also interested in the correlation between the Hot and Cold patterns observed in ZHENG gastritis patients and their corresponding symptom profiles.

Various types of features have been explored for tongue feature extraction and tongue analysis, including texture [5], color [6–8], shape [9], spectrum [8], among others. A systematic tongue feature set, comprising of a combination of geometric features (size, shape, etc.), cracks, and textures, was later proposed by Zhang et al. [10]. Computer-aided tongue analysis systems based on these types of features have also been developed [11, 12]. Our goal is to provide a set of objective features that can be extracted from patients’ tongue images, based on the knowledge of ZHENG, which improves accuracy of an objective clinical diagnosis. Our proposed tongue feature set is based on an extensive color model.

This paper is organized as follows: in Section 2, we provide a TCM descriptive view of the physiology of the tongue. An overview of the proposed feature extraction and learning framework along with a complete description of the color space model feature set is presented in Section 3. Our experimental results and analysis in a tongue image dataset from gastritis patients with Cold ZHENG and Hot ZHENG are discussed in Section 4 before drawing our conclusions and proposing plans for future work in Section 5.

2. Tongue Diagnosis in TCM

TCM believes that the tongue has many relationships and connections in the human body, both to the meridians and the internal organs. It is, therefore, very useful and important during inspection for confirming TCM diagnosis as it can present strong visual indicators of a person’s overall physical and mental harmony or disharmony. In TCM, the tongue is divided into tongue tip, tongue margins, tongue center, and tongue root. Figure 1(a) shows each part of the tongue and its correspondence to different internal organs according to TCM while Figure 1(b) illustrates how we geometrically obtain an approximation of these regions from the tongue image. The tongue tip reflects the pathological changes in the heart and lungs, while the bilateral sides of the tongue reflect that of the liver and gallbladders. The pathological changes in the spleen and stomach are mirrored by the center of tongue, while changes in the kidneys, intestines, and bladder section correspond to the tongue root.

In this paper, we focus on the patients with stomach disorders, gastritis. Hence, we are interested in extracting features not only from entire tongue image but also specifically from the middle region, as this corresponds to the stomach organ, according to TCM. We extract the middle rectangular region, illustrated in Figure 1(b), as our approximation for the tongue middle region.

The practitioner examines the general and local shape as well as the color of the tongue and its coating. According to TCM, the normal tongue is pale red with thin white coating. Some signs of imbalance or pathology are red body, yellow coating, or thick coating like mozzarella cheese, and so forth. Some characteristic changes occur in the tongue in some particular diseases. Most tongue attributes are on the tongue surface. A TCM doctor looks at several attributes of tongue body: color, moisture, size, shape, and coating. These signs not only reveal overall states of health but they also correlate to specific organ functions and disharmonies, especially in the digestive system.

The two main characteristics of the tongue in TCM ZHENG diagnosis are the color and the coating. The color of the patient’s tongue color provides information about his/her health status. For example [13], dark red color can indicate inflammation or ulceration, while a white tongue indicates cold attack, mucus deposits, or a weakness in the blood leading to such conditions as anemia [12]. Moreover, a yellow tongue points out a disorder of the liver and gallbladder, and blue or purple implies stagnation of blood circulation and a serious weakening of the part of the digestive system that corresponds to the area of the tongue where the color appears.

The coating on the tongue is discriminated by not only its presence but also its color. The color could be yellow, white, and other colors. However, the color in image is not the exact true color of the tongue. To properly identify the color of the tongue coating, we applied the specular component technique presented in our prior work on tongue detection and analysis [2]. Figure 2 illustrates different tongue images of patients and their corresponding ZHENG class.

3. Tongue Feature Extraction and Classification Framework

3.1. Feature Extraction for Tongue Image Analysis. Our goal is to compute a set of objective features \( \bar{F}_j = \{F_n\} \) from each tongue image \( j \) that can be fed into our learning
system so that we can predict not only the color and coating on the tongue, but also different ZHENGs of the gastritis patients. These features are designed to capture different color characteristics of the tongue. While a single feature may not be very discriminative, our premise is that the aggregation of these features will be discriminative. We leave it to the learning algorithm to determine the weight/contribution of each feature in the final classification.

Most color spaces are represented in tuples of number, normally three or four color components. Color components determine the position of the color in the color space used. There are many color spaces defined for different purposes. We designed a set of 25 features that span the entire color-space model. They can be grouped under eight categories: RGB, HSV, YIQ, Y’CbCr, XYZ, L*a*b*, CIE Luv, and CMYK.

In this section, we first describe in detail how we compute each feature \( f_{i} \) per \( i \)th pixel in the image. Then, we explain how each feature per pixel is aggregated to obtain \( F_{j} = \{ F_{n} \} \) per tongue image \( j \).

3.1.1. RGB. RGB is an additive color system, based on trichromatic theory in which red, green, and blue light components are added together to produce a specific pigment. The RGB model encodes the intensity of red, green, and blue, respectively. \((R_{i}, G_{i}, B_{i})\) for each pixel is an unsigned integer between 0 and 255. Each RGB feature \( \{ f_{n} \mid n = 1, \ldots, 3 \} \) represents the normalized intensity value of the red, green, and blue component, respectively, of the \( i \)th pixel in the image. We denote the normalized value of each component
as \( r_i = R_i/255 \), \( g_i = G_i/255 \), and \( b_i = B_i/255 \). Thus, \( f_{i1} = r_i; f_{i2} = g_i; f_{i3} = b_i \).

All the remaining color-space model features described in our feature set derive their value from the RGB feature set.

### 3.1.2. HSV

HSV color space represents color using a 3-tuple set of hue, saturation, and value. It separates the luminance component of the color from chrominance information. The HSV model \((H_i, S_i, V_i)\) is obtained by a linear transformation of normalized RGB color space \(\{r_i, g_i, b_i\}\).

For each pixel \(p_i\), let \(\tilde{M}_i = \max\{r_i, g_i, b_i\}\) represent the maximum value of the pixel's RGB triple set while \(\hat{m}_i = \min(r_i, g_i, b_i)\), the minimum value of the set. We also denote the difference between maximum and minimum values of each RGB tuple by \(\Delta_i = \tilde{M}_i - \hat{m}_i\). The HSV components \(\{H_i, S_i, V_i\}\) are computed from RGB color space \(\{r_i, g_i, b_i\}\) as follows:

\[
\begin{align*}
V_i &= \tilde{M}_i, \\
S_i &= \frac{\Delta_i}{M_i}, \quad \text{otherwise}, \\
H_i &= \begin{cases} \\
0, & \Delta_i = 0, \\
\frac{g_i - b_i}{b_i - r_i}, & \Delta_i > 0, \\
\frac{r_i - g_i}{\Delta_i + 2}, & \Delta_i < 0
\end{cases}
\end{align*}
\]

Thus, the HSV features are \(f_{i1} = H_i; f_{i2} = S_i; f_{i3} = V_i\).

### 3.1.3. YIQ

The YIQ color model is the television transmission color space for a digital standard. The Y component represents the perceived luminance, while I and Q components are the color information. I character is referred to “in-phase” term and Q letter stands for “quadrature.” I and Q can place color in a graph representing I as X axis and Q as Y axis. The YIQ system takes advantage of human color perceiver characteristics [14, 15].

The YIQ model \((Y_i, I_i, Q_i)\) is obtained by a linear transformation of the normalized RGB color space \(\{r_i, g_i, b_i\}\) as follows:

\[
\begin{bmatrix}
Y_i \\
I_i \\
Q_i
\end{bmatrix} = \begin{bmatrix}
0.299 & +0.587 & +0.114 \\
0.596 & -0.274 & -0.222 \\
0.212 & -0.539 & +0.532
\end{bmatrix} \begin{bmatrix}
r_i \\
g_i \\
b_i
\end{bmatrix}.
\]

The \(\{Y_i, I_i, Q_i\}\) values are each normalized to obtain \(\{Y_i', I_i', Q_i\} \in [0, 1]\). Thus, the YIQ features are \(f_{i1} = Y_i'; f_{i2} = I_i'; f_{i3} = Q_i\).

### 3.1.4. Y’CbCr

Like YIQ, Y’CbCr is the television transmission color space but it is in analogue spaces for the NTSC system. Y’CbCr color space detaches RGB into the luma component, the blue-difference and red-difference chroma components. The transformation equation from RGB (unnormalized) model to YCbCr is defined as

\[
\begin{bmatrix}
Y_i' \\
C_{bi} \\
C_{ri}
\end{bmatrix} = \begin{bmatrix}
0.299 & +0.587 & +0.114 \\
-0.169 & -0.331 & +0.500 \\
0.500 & -0.419 & -0.081
\end{bmatrix} \begin{bmatrix}
r_i \\
g_i \\
b_i
\end{bmatrix}.
\]

Similar to the YIQ features, \(\{Y_i', C_{ri}, C_{bi}\}\) values are each normalized to obtain \(\{y_i', cb_i, cr_i\} \in [0, 1]\). Thus the YIQ features are \(f_{i1} = y_i'; f_{i2} = cb_i; f_{i3} = cr_i\).

### 3.1.5. XYZ

Brightness and chromaticity are two principal components of color that interact with human vision. XYZ are developed under CIE XYZ color space [16]. The XYZ values can be obtained by a linear transformation of the gamma corrected value of the RGB normalized color space \(\{r_i, g_i, b_i\}\).

The gamma-corrected function is defined as

\[
y(t) = \begin{cases} \\
\frac{t}{12.92} & \text{if } t \leq 0.04045, \\
\left(\frac{t}{1 + a}\right)^{2.4} & \text{otherwise},
\end{cases}
\]

where \(a = 0.055\). Thus, XYZ model consisting of \(\{X_i', Y_i', Z_i\}\) components is given by

\[
\begin{bmatrix}
X_i' \\
Y_i' \\
Z_i
\end{bmatrix} = \begin{bmatrix}
0.4124 & 0.3576 & 0.1805 \\
0.2126 & 0.7152 & 0.0722 \\
0.0193 & 0.1192 & 0.9505
\end{bmatrix} \begin{bmatrix}
y(r_i) \\
y(g_i) \\
y(b_i)
\end{bmatrix}.
\]

The \(\{X_i', Y_i', Z_i\}\) values are each normalized to obtain \(\{x_i, y_i', z_i\} \in [0, 1]\). Thus, the XYZ features are defined as \(f_{i1} = x_i; f_{i2} = y_i'; f_{i3} = z_i\).

### 3.1.6. L*a*b*

CIE L*a*b* color space is a nonlinear transformation of the CIE XYZ color space [17]. CIE L*a*b* try to imitate the logarithmic response of the human eye. The L* component is designed to match closely with human perception of lightness. The other two components describe the chroma.

The forward transformation of CIE XYZ color space to CIE L*a*b* is computed as follows:

\[
\begin{align*}
L_i^* &= 116\varphi\left(\frac{Y_i''}{\delta_2}\right) - 16, \\
A_i &= 500\left[\varphi\left(\frac{X_i}{\delta_1}\right) - \varphi\left(\frac{Y_i''}{\delta_2}\right)\right], \\
B_i &= 200\left[\varphi\left(\frac{Y_i''}{\delta_2}\right) - \varphi\left(\frac{Z_i}{\delta_3}\right)\right],
\end{align*}
\]

where

\[
\varphi(t) = \begin{cases} \\
\frac{t^{1/3}}{\delta_3} & \text{if } t > \left(\frac{6}{29}\right)^3, \\
\frac{1}{\delta_3}\left(\frac{29}{6}\right)\left(t + \frac{4}{29}\right) & \text{otherwise},
\end{cases}
\]

and \(\delta\) denotes the D65 white point given by \(\{0.950456, 1.0, 1.088754\}\).

The L*a*b* values \(\{L_i^*, A_i, B_i\}\) are normalized as \(\{l_i^*, a_i, b_i\} \in [0, 1]\). Hence, the CIE L*a*b* color features are given by \(f_{i6} = l_i^*; f_{i7} = a_i; f_{i8} = b_i\).
3.1.7. CIE Luv. CIE Luv, or L* u* v*, is color-space- computed from the transformation of the CIE XYZ color space by International Commission on Illumination (CIE) in order to perceptual uniformity [17]. Similar to CIE L* a* b*, the D65 white point is referred by {δ}:

\[
L_i'' = \begin{cases} 
\left( \frac{29}{3} \right)^3 \left( \frac{Y_i''}{\delta_2} \right), & \text{if } \frac{Y_i''}{\delta_2} \leq \left( \frac{6}{29} \right)^3, \\
116 \left( \frac{Y_i''}{\delta_2}^{1/3} - 16 \right), & \text{otherwise,}
\end{cases}
\]

where \( k_1 = 0.2009, k_2 = 0.4610, \) under the standard luminance C. The normalized \( \{L_i'', U_i, V_i\} \) values are denoted by \( \{l_i'', u_i, v_i\} \in [0, 1] \). Therefore, \( f_{y0} = l_i''; f_{y0} = u_i; f_{y1} = v_i \).

3.1.8. CMYK. The CMYK color space is a subtractive color system mainly used in the printing industry [16]. The components consist of cyan, magenta, yellow, and neutral black. It is a common way to translate RGB display on monitors to CMYK values for printing.

Let \( \bar{M}_i = \max \{r_i, g_i, b_i\} \) represent the maximum value of the pixel’s RGB triple set. The CMYK color space, denoted by \( \{C_i, M_i, Y_i, K_i\} \), can be computed from the RGB model as follows:

\[
\begin{align*}
K_i &= 1 - \bar{M}_i, \\
C_i &= \frac{\bar{M}_i - r_i}{\bar{M}_i}, \\
M_i &= \frac{\bar{M}_i - g_i}{\bar{M}_i}, \\
Y_i &= \frac{\bar{M}_i - b_i}{\bar{M}_i},
\end{align*}
\]

Thus, the CMYK features are computed as \( f_{23} = C_i; f_{23} = M_i; f_{25} = Y_i; f_{25} = K_i \).

3.1.9. Aggregate Operators for the Feature Vectors. To train our classification model using this set of features, we need to combine the features per pixel into one composite feature vector \( \bar{F}_j = \{F_n\} \) per tongue image (or region) \( j \). We aggregate the pixel features using two different statistical averages (mean and median) and the standard deviation values. We derive five variations of feature vectors for our automated tongue ZHENG classification system using the following operators: mean, median (\( \text{med}\bar{F} \)), standard deviation (\( \sigma\bar{F} \)), “mean plus standard deviation” (\( \{\mu\bar{F}, \sigma\bar{F}\} \)), and “median plus standard deviation” (\( \{\text{med}\bar{F}, \sigma\bar{F}\} \)). Let \( N \) denote the number of pixels in a given tongue image (or region) \( j \). The mean feature vector is denoted by \( \mu\bar{F}_j = \{\mu F_n\} \), where \( \mu F_n \) is given by

\[
\mu F_n = \frac{\sum_{i=1}^{N} f_{ni}}{N}, \quad n = 1, \ldots, 25.
\]

The median feature vector, denoted by \( \text{med}\bar{F}_j = \{\text{med} F_n\} \), is computed as \( \text{med} F_n = \text{mid} \{\text{sort}(F_{ni})\}, \) \( n = 1, \ldots, 25 \). Standard deviation depicts the margin of difference between a given feature value and its average value among all the pixels in the given region. Thus, the standard deviation feature vector is denoted by \( \sigma\bar{F}_j = \{\sigma F_n\} \), where \( \sigma F_n \) is given by

\[
\sigma F_n = \frac{\sqrt{\sum_{i=1}^{N}(f_{ni} - \mu F_n)^2}}{N}, \quad n = 1, \ldots, 25.
\]

The “mean plus standard deviation,” denoted by \( \{\mu\bar{F}, \sigma\bar{F}\} \), is a concatenation of the mean feature vector and the standard deviation feature vector. Similarly, the “median plus standard deviation” feature vector, denoted by \( \{\text{med}\bar{F}, \sigma\bar{F}\} \), is a concatenation of the median feature vector and the standard deviation feature vector. Thus, the total number of features in both concatenated feature vectors is 50 each.

3.2. Supervised Learning Algorithms for ZHENG Classification. We apply three different supervised learning algorithms (AdaBoost, support vector machine, multilayer perceptron network) to build classification models for training and evaluating the proposed automated tongue based diagnosis system. Each model has its strength and weakness, which we describe briefly below. We empirically evaluate their performance over our dataset.

3.2.1. AdaBoost. An ensemble of classifiers is a set of classifiers whose individual predictions are combined in some way (typically by voting) to classify new examples. Boosting is a type of ensemble classifier which generates a set of weak classifiers using instances drawn from an iteratively updated distribution of the data, where in each iteration the probability of incorrectly classified examples is increased and the probability of the correctly classified examples is decreased. The ensemble classifier is a weighted majority vote of the sequence of classifiers produced.

The AdaBoost algorithm [18] trains a weak or base-learning algorithm repeatedly in a series of round \( t = 1, \ldots, T \). Given a training set \( \{x_i, y_i\}_{i=1}^{n} \), where \( x_i \) belongs to some domain \( X \) and \( y_i \in Y = \{-1, +1\} \) (the corresponding binary class labels), we denote the weight of \( i \)th example in round \( t \) by \( D_t(i) \). Initially, all weights are set equally and so \( D_t(i) = 1/n \) for all \( i \). For each round \( t \), a weak learner is trained using the current distribution \( D_t \). When we obtain a weak hypothesis \( h_t \) with error \( e_t =...
of the gastritis patients have been classified as Hot or Cold ZHENG and are identified with a color label (yellow or white) based on the color of the coating of their tongue, as determined by their Chinese doctors. The doctors also carry out a detailed profile of the ZHENG symptoms for each patient based on clinical evaluations. The list of the main symptom profile terms is summarized in Table 1.

We are also interested in the relationship between TCM diagnosis and Western medicine diagnosis; hence, for a subset of the patients, we are provided with their corresponding Western medical gastritis pathology. They are grouped into two categories: superficial versus atrophic. In Western medicine, the doctors are also interested in knowing whether the Helicobacter Pylori (HP) bacterium found in the stomach is present (positive) or absent (negative) in the patients with chronic gastritis. Thus, we are provided with that information for a subset of the patients. It was not feasible to obtain all the different information collected per patient. Table 2 summarizes the population of each subset for four different labels (ZHENG, Coating, Pathology, and HP).

4. Results and Analysis

4.1. Experimental Setup. In this section, we evaluated the performance of our proposed ZHENG classification system using the three classification models (AdaBoost, SVM, and MLP) described in Section 3.2. We compared the performance of training the classifier models using the set of features extracted from the entire tongue image versus the middle tongue region only. As mentioned in Section 2, in TCM, it is believed that the middle tongue region provides discriminant information for diagnosing stomach disorders. Hence, we extract features from the middle tongue region, as described in Figure 1(b), to evaluate the performance compared to extracting features from the entire tongue region. In training and testing our classification models, we employ a 3-fold cross-validation strategy. This implies that the data is split into three sets; one set is used for testing and the remaining two sets are used for training. The experiment is repeated with each of the three sets used for testing. The average accuracy of the tests over the three sets is taken as the performance measure. For each classification model, we varied the parameters to optimize its performance. We also compare the results obtained using the five different variations of the feature vector (mean = μF, median = med F, standard deviation = σF, mean + standard deviation = {μF, σF}, and median + standard deviation = {med F, σF}), as described in Section 3.1. We also apply Information Gain attribute evaluation on the feature vectors to quantify and rank the significance of individual features. Lastly, we apply the Best First feature selection algorithm to select the “significant” features before training the classifiers to compare the performance of training the classifiers with the whole feature set against selected features.

The performance metrics used are the classification accuracy (CA) and the average F-measure. CA is defined as the percentage of correctly classified instances over the entire set of instances classified. In our dataset, as described
The experimental results show that our proposed color-space feature vector can automatically classify the color of the coating of a gastritis patient’s tongue as yellow or white. We can observe from Table 3 that the combination of the median and standard deviation feature values \( \{ \text{med} \vec{F}, \sigma \vec{F} \} \) yields the best result for both the entire tongue region and the middle tongue region only. The results for both regions are also very comparable.

When using the entire tongue region, the top three significant features for the color coating classification, ranked by the information gain attribute, were \( \{ \sigma F_9, \text{med} F_{12}, \sigma F_1 \} \), which denote the standard deviation of \( Q \) chroma (YIQ model), the median of \( C_r \) component (YCbCr), and the standard deviation of Green Channel (RGB), respectively. For the middle tongue region only, the top three were \( \{ \sigma F_9, \sigma F_{20}, \text{med} F_4 \} \) which denote the standard deviation of \( Q \) chroma (YIQ model), the standard deviation of \( u \) component (\( L^*u^*v^* \)), and the median of the Hue (HSV) component. It is also interesting to observe that out of the top ten significant features using the entire region versus the middle tongue region, they both have six of those features in common.

The result obtained on ZHENG classification between the Hot and Cold groups is shown in Table 4. For the ZHENG classification, using the standard deviation feature values \( \{ \sigma \vec{F} \} \) performs best when dealing with the entire tongue region while the \( \{ \text{med} \vec{F}, \sigma \vec{F} \} \) feature vector is the top performer for the middle tongue region only.

For ZHENG classification between Hot and Cold syndromes for gastritis patients, when using the entire tongue region, only one feature was considered significant by the information gain attribute: \( \sigma F_9 \), that is, which is the standard deviation of \( Q \) chroma (YIQ model). For the middle tongue region, the most important feature is \( \sigma F_{20} \), the standard deviation of \( u \) component (\( L^*u^*v^* \)). Even though the noteworthy feature in the entire tongue area and the middle tongue area is not the same, both \( Q \) components in YIQ color space and \( u \) component in \( L^*u^*v^* \) color space show the difference from green to red in chromaticity diagram.

Table 5 summarizes the results obtained when we train different classifiers to detect the presence of the HP bacteria in a gastritis patient using the color feature vector. The classification result obtained in learning the pathology groups of the patients (superficial versus atrophic) is shown in Table 6. Both cases are not very strong, which illustrates a weak correlation between the western medicine diagnosis and the tongue information utilized by Chinese medical practitioners. No feature was identified as significant in either case.
<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu F$</td>
<td>0.618</td>
<td>0.659</td>
</tr>
<tr>
<td>$(\mu F, \sigma F)$</td>
<td>0.644</td>
<td>0.669</td>
</tr>
<tr>
<td>$\text{med} F$</td>
<td>0.655</td>
<td>0.669</td>
</tr>
<tr>
<td>$(\text{med} F, \sigma F)$</td>
<td>0.661</td>
<td>0.695</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.763</td>
<td>0.763</td>
</tr>
</tbody>
</table>

Table 4: ZHENG classification between Hot and Cold syndromes for gastritis patients.

<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu F$</td>
<td>0.679</td>
<td>0.695</td>
</tr>
<tr>
<td>$(\mu F, \sigma F)$</td>
<td>0.702</td>
<td>0.702</td>
</tr>
<tr>
<td>$\text{med} F$</td>
<td>0.697</td>
<td>0.697</td>
</tr>
<tr>
<td>$(\text{med} F, \sigma F)$</td>
<td>0.677</td>
<td>0.677</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.763</td>
<td>0.763</td>
</tr>
</tbody>
</table>

Table 5: Detection of presence of HP bacteria (positive versus negative) in gastritis patients.

<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu F$</td>
<td>0.604</td>
<td>0.647</td>
</tr>
<tr>
<td>$(\mu F, \sigma F)$</td>
<td>0.633</td>
<td>0.661</td>
</tr>
<tr>
<td>$\text{med} F$</td>
<td>0.633</td>
<td>0.695</td>
</tr>
<tr>
<td>$(\text{med} F, \sigma F)$</td>
<td>0.657</td>
<td>0.697</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.637</td>
<td>0.697</td>
</tr>
</tbody>
</table>

Table 6: Classification between superficial and atrophic pathology of the gastritis patients.

<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu F$</td>
<td>0.579</td>
<td>0.659</td>
</tr>
<tr>
<td>$(\mu F, \sigma F)$</td>
<td>0.676</td>
<td>0.676</td>
</tr>
<tr>
<td>$\text{med} F$</td>
<td>0.687</td>
<td>0.702</td>
</tr>
<tr>
<td>$(\text{med} F, \sigma F)$</td>
<td>0.717</td>
<td>0.717</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.701</td>
<td>0.761</td>
</tr>
</tbody>
</table>

Table 7: Tongue Classification between superficial and atrophic in Cold syndrome patients.
<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu F$</td>
<td>0.768</td>
<td>77.06</td>
<td>0.735</td>
<td>73.39</td>
<td>0.710</td>
<td>71.56</td>
</tr>
<tr>
<td>$(\mu F, \sigma F)$</td>
<td>0.741</td>
<td>74.31</td>
<td>0.845</td>
<td>84.40</td>
<td>0.764</td>
<td>76.15</td>
</tr>
<tr>
<td>$\text{med} F$</td>
<td>0.718</td>
<td>72.48</td>
<td>0.708</td>
<td>70.09</td>
<td>0.690</td>
<td>69.35</td>
</tr>
<tr>
<td>$(\text{med} F, \sigma F)$</td>
<td>0.715</td>
<td>71.36</td>
<td>0.817</td>
<td>81.65</td>
<td>0.672</td>
<td>67.89</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.770</td>
<td>77.06</td>
<td>0.818</td>
<td>81.65</td>
<td>0.675</td>
<td>67.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu F$</td>
<td>0.583</td>
<td>59.68</td>
<td>0.773</td>
<td>77.42</td>
<td>0.705</td>
<td>70.97</td>
</tr>
<tr>
<td>$(\mu F, \sigma F)$</td>
<td>0.740</td>
<td>74.19</td>
<td>0.839</td>
<td>83.87</td>
<td>0.765</td>
<td>77.42</td>
</tr>
<tr>
<td>$\text{med} F$</td>
<td>0.628</td>
<td>62.90</td>
<td>0.740</td>
<td>74.19</td>
<td>0.675</td>
<td>67.74</td>
</tr>
<tr>
<td>$(\text{med} F, \sigma F)$</td>
<td>0.774</td>
<td>77.42</td>
<td>0.839</td>
<td>83.87</td>
<td>0.757</td>
<td>75.81</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.834</td>
<td>83.87</td>
<td>0.757</td>
<td>83.87</td>
<td>0.839</td>
<td>83.87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu F$</td>
<td>0.539</td>
<td>55.14</td>
<td>0.642</td>
<td>63.55</td>
<td>0.645</td>
<td>64.49</td>
</tr>
<tr>
<td>$(\mu F, \sigma F)$</td>
<td>0.662</td>
<td>67.29</td>
<td>0.681</td>
<td>69.16</td>
<td>0.698</td>
<td>70.09</td>
</tr>
<tr>
<td>$\text{med} F$</td>
<td>0.612</td>
<td>61.68</td>
<td>0.646</td>
<td>63.55</td>
<td>0.666</td>
<td>66.36</td>
</tr>
<tr>
<td>$(\text{med} F, \sigma F)$</td>
<td>0.704</td>
<td>71.03</td>
<td>0.657</td>
<td>64.49</td>
<td>0.677</td>
<td>68.22</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.834</td>
<td>83.87</td>
<td>0.757</td>
<td>75.81</td>
<td>0.838</td>
<td>83.87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu F$</td>
<td>0.803</td>
<td>82.82</td>
<td>0.831</td>
<td>82.44</td>
<td>0.795</td>
<td>80.53</td>
</tr>
<tr>
<td>$(\mu F, \sigma F)$</td>
<td>0.829</td>
<td>83.59</td>
<td>0.851</td>
<td>85.11</td>
<td>0.848</td>
<td>85.50</td>
</tr>
<tr>
<td>$\text{med} F$</td>
<td>0.785</td>
<td>80.53</td>
<td>0.803</td>
<td>83.21</td>
<td>0.814</td>
<td>83.21</td>
</tr>
<tr>
<td>$(\text{med} F, \sigma F)$</td>
<td>0.814</td>
<td>83.21</td>
<td>0.835</td>
<td>83.59</td>
<td>0.861</td>
<td>86.26</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.818</td>
<td>83.21</td>
<td>0.839</td>
<td>83.59</td>
<td>0.851</td>
<td>85.11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
<th>Entire tongue</th>
<th>Middle tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu F$</td>
<td>0.765</td>
<td>78.63</td>
<td>0.809</td>
<td>80.24</td>
<td>0.784</td>
<td>78.63</td>
</tr>
<tr>
<td>$(\mu F, \sigma F)$</td>
<td>0.836</td>
<td>84.68</td>
<td>0.852</td>
<td>84.68</td>
<td>0.857</td>
<td>85.89</td>
</tr>
<tr>
<td>$\text{med} F$</td>
<td>0.756</td>
<td>77.82</td>
<td>0.795</td>
<td>81.45</td>
<td>0.784</td>
<td>78.63</td>
</tr>
<tr>
<td>$(\text{med} F, \sigma F)$</td>
<td>0.802</td>
<td>81.45</td>
<td>0.845</td>
<td>84.27</td>
<td>0.844</td>
<td>84.68</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.826</td>
<td>83.47</td>
<td>0.849</td>
<td>84.68</td>
<td>0.843</td>
<td>84.27</td>
</tr>
</tbody>
</table>
The next set of experimental results focus on training our classifier using our proposed color-space feature vector to discriminate Hot ZHENG from Cold ZHENG in each pathology group. Table 9 summarizes the results obtained to train the classifiers to automatically classify between Hot and Cold ZHENG for superficial gastritis patients. Table 10 reflects the results for gastritis patients. We can observe from Table 9 that both $\{\mu F, \sigma F\}$ and $\{\text{med} F, \sigma F\}$ feature vectors perform the best for both the entire tongue region and the middle tongue region. From results in Table 10, using the standard deviation feature values ($\{\mu F, \sigma F\}$) performs best when dealing with the entire tongue region while the $\{\text{med} F, \sigma F\}$ feature vector is the top performer for the middle tongue region.

When using the entire tongue region, the top three significant features for the pathology classification between superficial and atrophic in Hot syndrome, ranked by the information gain attribute, were $\{\sigma F_9, \sigma F_6, \sigma F_1\}$ which denote the standard deviation of $Q$ chroma (YIQ model), the standard deviation of value component (HSV), and the standard deviation of Red Channel (RGB), respectively.

In Table 8, when using the entire tongue region, the top three significant features for the pathology classification between superficial and atrophic in Hot syndrome, ranked by the information gain attribute, were $\{\mu F_{23}, \mu F_{25}, \mu F_5\}$ which denote the mean of Cyan Ink (CMYK model), the mean of Black Ink (CMYK model), and the mean of Blue Channel (RGB), respectively. For the middle tongue region only, the top three were $\{\sigma F_{22}, \sigma F_{25}, \text{med} F_{25}\}$, which denote the standard deviation of Cyan Ink (CMYK model), the standard deviation of Black Ink (CMYK model), and the median of Black Ink (CMYK model).

When using the entire tongue region, the top three significant features for the ZHENG classification between Hot syndrome and Cold syndrome in the patients who are superficial, ranked by the information gain attribute, were $\{\sigma F_9, \text{med} F_5, \text{med} F_{18}\}$, which denote the standard deviation of $Q$ chroma (YIQ model), the median of Blue Channel (RGB), and the median of the blue sensitivity $Z$ component, respectively. For the middle tongue region only, the top three were $\text{med} F_{24}, \sigma F_{19},$ and $\text{med} F_5$ which denote

### Table 13: Tongue classification between normal group and Hot ZHENG.

<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th></th>
<th></th>
<th>Middle tongue</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AdaBoost</td>
<td>SVM</td>
<td>MLP</td>
<td>AdaBoost</td>
<td>SVM</td>
<td>MLP</td>
</tr>
<tr>
<td>$\mu F$</td>
<td>0.671</td>
<td>70.00</td>
<td>0.781</td>
<td>77.78</td>
<td>0.708</td>
<td>72.22</td>
</tr>
<tr>
<td>${\mu F, \sigma F}$</td>
<td>0.804</td>
<td>80.56</td>
<td>0.792</td>
<td>79.44</td>
<td>0.816</td>
<td>81.67</td>
</tr>
<tr>
<td>med $F$</td>
<td>0.721</td>
<td>72.78</td>
<td>0.711</td>
<td>72.22</td>
<td>0.739</td>
<td>75.00</td>
</tr>
<tr>
<td>${\text{med} F, \sigma F}$</td>
<td>0.796</td>
<td>80.00</td>
<td>0.814</td>
<td>82.78</td>
<td>0.797</td>
<td>80.00</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.768</td>
<td>77.22</td>
<td>0.828</td>
<td>82.22</td>
<td>0.826</td>
<td>82.78</td>
</tr>
</tbody>
</table>

### Table 14: Tongue classification between normal group and Cold ZHENG.

<table>
<thead>
<tr>
<th>Feature vector</th>
<th>Entire tongue</th>
<th></th>
<th></th>
<th>Middle tongue</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AdaBoost</td>
<td>SVM</td>
<td>MLP</td>
<td>AdaBoost</td>
<td>SVM</td>
<td>MLP</td>
</tr>
<tr>
<td>$\mu F$</td>
<td>0.690</td>
<td>68.97</td>
<td>0.759</td>
<td>75.86</td>
<td>0.676</td>
<td>68.10</td>
</tr>
<tr>
<td>${\mu F, \sigma F}$</td>
<td>0.742</td>
<td>74.14</td>
<td>0.785</td>
<td>78.45</td>
<td>0.748</td>
<td>75.00</td>
</tr>
<tr>
<td>med $F$</td>
<td>0.686</td>
<td>68.97</td>
<td>0.745</td>
<td>75.00</td>
<td>0.757</td>
<td>75.86</td>
</tr>
<tr>
<td>${\text{med} F, \sigma F}$</td>
<td>0.759</td>
<td>75.86</td>
<td>0.774</td>
<td>77.59</td>
<td>0.734</td>
<td>73.28</td>
</tr>
<tr>
<td>$\sigma F$</td>
<td>0.741</td>
<td>74.14</td>
<td>0.733</td>
<td>73.28</td>
<td>0.734</td>
<td>73.28</td>
</tr>
</tbody>
</table>
and tongue with coating, ranked by the information gain significant features for distinguishing between normal tongue feature vector. While for the middle tongue region, it was the which denotes the standard deviation of

When using the entire tongue region, there is only one significant feature difference for the ZHENG classification between Hot syndrome and Cold syndrome in patients who are atrophic, ranked by the information gain attribute, \( \sigma F_3 \) which denotes the standard deviation of \( Q \) chroma (YIQ model). For the middle tongue region only, there were two significant features: \( \{ \mu F_{19}, \mu F_3 \} \) which denote the mean of the blue sensitivity \( Z \) component (XYZ) and the mean of the Blue Channel (RGB).

4.3. Classification Results for Gastritis Patients versus Control Group. The experimental results presented in this section analyze the discrimination between the gastritis patients and control group. Table 11 summarizes the results obtained using our proposed color-space feature vector to train the classifiers to automatically classify patients with coating on tongue versus healthy patients with normal tongue (without coating). We can observe from Table 11 that the \( \{ \text{med} \bar{F}, \sigma \bar{F} \} \) feature vector yields the best result for the entire tongue region while for the middle tongue region, it was the \( \sigma \bar{F} \) feature vector.

When using the entire tongue region, the top three significant features for distinguishing between normal tongue and tongue with coating, ranked by the information gain attribute, were \( \{ \sigma F_1, \sigma F_6, \sigma F_{25} \} \) which denote the standard deviation of Red Channel (RGB), the standard deviation of value component (HSV), and the standard deviation of Black Ink (CMYK) respectively. For the middle tongue region only, there were only two significant features: \( \{ \sigma F_{13}, \sigma F_{14} \} \) which denote the standard deviation of lightness component (L∗a+b∗) and the standard deviation of \( a^* \) component (L∗a+b∗). It is also interesting to observe that by comparing the set of the top 10 significant features using the entire region versus the set from the middle tongue region, they both have the lightness and \( a^* \) component (L∗a+b∗) in common.

The results obtained from the classification between the normal group and the entire set of patients with ZHENG syndrome is shown in Table 12. The \( \{ \mu \bar{F}, \sigma \bar{F} \} \) feature vector performs best when dealing with the entire tongue region while the \( \{ \text{med} \bar{F}, \sigma \bar{F} \} \) feature vector is the top performer for the middle tongue region.

4.4. Classification Results for Gastritis Patients versus Hot ZHENG Group. The experimental results presented in this section analyze the discrimination between the gastritis patients and Hot ZHENG patients only, and then normal group versus Cold ZHENG patients only. Table 13 illustrates the results for normal versus hot ZHENG. We can observe that the \( \sigma \bar{F} \) feature vector performs best both for the entire tongue region and the middle tongue region. From Table 14, when only the normal versus Cold ZHENG patients is considered,
the same feature vector, \(\{\mu_{\vec{F}}, \sigma_{\vec{F}}\}\), performs best for both cases, however, considering only the middle tongue region outperforms using the entire tongue region.

When using the entire tongue region, the top three significant features for the classification between the normal group and the gastritis patients with Hot syndrome, ranked by the information gain attribute, were \(\{\sigma F_1, \sigma F_6, \sigma F_{25}\}\) which denote the standard deviation of Red Channel (RGB), the standard deviation of value component (HSV), and the standard deviation of Black Ink (CMYK), respectively. For the middle tongue region only, there were only two significant features: \(\{\sigma F_{13}, \sigma F_{14}\}\) which denote the standard deviation of lightness component \((L^*a^*b^*)\) and the standard deviation of \(a^*\) component \((L^*a^*b^*)\). When the set of the top ten significant features using the entire region versus the set from the middle tongue region are compared, they both have the lightness and \(a^*\) component \((L^*a^*b^*)\) in common.

When using the entire tongue region, the top three significant features for the classification between the normal group and the gastritis patients with Cold syndrome, ranked by the information gain attribute, were \(\{\sigma F_{25}, \sigma F_{22}, \sigma F_1\}\) which denote the standard deviation of Black Ink (CMYK), the standard deviation of Cyan Ink (CMYK), and the standard deviation of Red Channel (RGB), respectively. For the middle tongue region only, the top three were \(\{\sigma F_{13}, \mu F_{22}, \sigma F_{14}\}\) which denote the standard deviation of lightness component \((L^*a^*b^*)\), the mean of Cyan Ink (CMYK), and the standard deviation of \(a^*\) component \((L^*a^*b^*)\).

Table 15 show the results of training our classifiers to discriminate between the normal group and the superficial patients while Table 16 shows the result for normal group versus the atrophic patients. When using the entire tongue region, the top three significant features for the classification between the normal group and the superficial group, ranked by the information gain attribute, were \(\{\sigma F_1, \sigma F_6, \sigma F_{25}\}\) which denote the standard deviation of Red Channel (RGB), the standard deviation of value component (HSV), and the standard deviation of Black Ink (CMYK), respectively. For the middle tongue region, the top three were \(\{\text{med} F_6, \text{med} F_{1}, \text{med} F_{6}\}\) which denote the median of \(Q\) chromatic component \((YIQ)\), the median of Red Channel (RGB), and the median of Value component (HSV).

When using the entire tongue region, the top three significant features for the classification between the normal group and the atrophic group, ranked by the information gain attribute, were \(\{\mu F_{25}, \mu F_{22}, \mu F_{1}\}\) which denote the mean of Black Ink (CMYK model), the mean of Cyan Ink (CMYK model), and the mean of Red Channel (RGB), respectively. For the middle tongue region, the top three were \(\{\text{med} F_{16}, \sigma F_{13}, \sigma F_{23}\}\) which denote the median of red sensitivity \(X\) component \((XYZ)\), the standard deviation of lightness \((L^*a^*b^*)\), and the standard deviation of Cyan Ink (CMYK).

4.4. Analysis of Classification Results. From the experimental results presented in Sections 4.2 and 4.3, we can draw the following conclusions. Firstly, concerning the performance of the different classification models, we observe that the MLP and SVM models usually outperformed the AdaBoost model. The multilayer perceptron neural network seems most adequate for learning the complex relationships between the color features of the tongue images and the ZHENG/coating classes. However, both the MLP and SVM models have many parameters to consider and optimize while the AdaBoost is a much simpler model. In the AdaBoost model, we use a decision tree as our base weak learner and vary the number of classifiers to optimize its performance.

Secondly, we observe that when making discriminations within the gastritis patients group (hot versus cold ZHENG, yellow versus white coating, etc.), it was more profitable to apply the feature vectors on the entire tongue image. When classifying the normal groups versus the ZHENG groupings, usually, it improved classifier performance to apply the feature vectors to the middle tongue regions only.
Thirdly, we also observe that from the evaluation of the variations of the feature vectors used, taking into account both the average and the standard deviation usually resulted in an excellent performance. It seemed like the mean outperformed the median slightly, overall, that is, \( \{ \mu \vec{F}, \sigma \vec{F} \} \). In a few cases, simply considering variation of the spread of the values over the region (\( \{ \sigma \vec{F} \} \)) yielded the best performance. Thus, we can conclude that when deriving a feature vector for the tongue image, the mean (or median) as well as the standard deviation (which takes into account the variation of the spread on the region) is very important.

Lastly, we observe that though we were not able to effectively discriminate between the pathology groups (superficial versus atrophic and also the presence of the HP bacterium using our color-space feature vectors, we were able to classify them much better when we took into account the ZHENG classes. This further strengthens the notion that our proposed color-space feature vectors are able to discriminate between the hot and cold ZHENG patients in addition to discerning a ZHENG patient from a non-ZHENG (healthy) patient.

4.5. Applying Feature Selection Algorithm. The classification results presented in Sections 4.2 and 4.3 were obtained using the whole feature set. For each experiment carried out on the entire tongue region, we also applied information gain attribute evaluation to rank the significance of the features. In this section, we apply feature selection algorithm (Best First) to select only a subset of features, which are deemed significant, before training the classifiers. Our goal is to see if this would yield a better result than using the whole feature set. The Best First algorithm searches the space of attribute subsets by greedy hill climbing augmented with a backtracking facility.

The summary of the results obtained is shown in Table 17. The normal group refers to the healthy (non-ZHENG) control group. We present the best classification result obtained for each experiment based on using the five variations of the feature vectors (\( \mu \vec{F}, \text{med} \vec{F}, \sigma \vec{F}, \{ \mu \vec{F}, \sigma \vec{F} \}, \{ \text{med} \vec{F}, \sigma \vec{F} \} \)) and the three different classification models (Adaboost, SVM, and MLP). As we can observe from Table 17, using the whole feature set to train the classifiers yielded a better result in all cases except for the Atrophic Patients (Hot versus Cold ZHENG) experiment. Thus, we can conclude the overall, using the aggregate of the proposed feature sets is more discriminative even though some features are more significant than others.

5. Conclusion and Future Work

In this paper, we propose a novel color space-based feature set for use in the clinical characterization of ZHENG using various supervised machine-learning algorithms. Using an automated tongue-image diagnosis system, we extract these objective features from tongue images of clinical patients and analyze the relationship with their corresponding ZHENG data and disease prognosis (specifically gastritis) obtained from clinical practitioners. Given that TCM practitioners usually observe the tongue color and coating to determine ZHENG (such as Cold or Hot ZHENG) and to diagnose different stomach disorders including gastritis. We propose using machine-learning techniques to establish the relationship between the tongue image features and ZHENG by learning through examples.

The experimental results obtained demonstrate an excellent performance of our proposed system. Our future work will focus on improving the performance of our system by exploring additional tongue image features that can be extracted to further strengthen our classification models. We plan to explore ways to improve our methodology to more accurately classify the ZHENGs such as including a preprocessing step of coating separation prior to the feature extraction phase. Lastly, we plan to evaluate the classification of the other ZHENG types mentioned in Section 1.

Acknowledgment

S. Li and T. Ma are supported in part by the NSFC (no. 90709013).

References


Submit your manuscripts at http://www.hindawi.com