ABSTRAK

Perbedaan tingkat fermentasi menghasilkan variasi karakter mutu yang mencakup perubahan senyawa bioaktif, senyawa volatil, rasa, dan aroma pada teh hijau (non-fermentasi), kuning (fermentasi ringan), merah (semi-fermentasi), dan hitam (fermentasi penuh), sehingga klasifikasi yang akurat menjadi penting untuk mendukung pengendalian mutu dan peningkatan daya saing produk. Aroma, yang merefleksikan profil senyawa volatil, menjadi fokus penelitian ini dan dianalisis menggunakan *electronic nose (e-nose)*. Dalam sistem *e-nose*, data yang dihasilkan berdimensi tinggi, sehingga pemilihan kombinasi fitur yang tepat dapat meningkatkan performa klasifikasi. Penelitian ini bertujuan mengembangkan metode klasifikasi teh berbasis *e-nose* dengan optimasi seleksi fitur berbasis skor *silhouette* di ruang *linear discriminant analysis* (LDA). Data *e-nose* dianalisis menggunakan lima metode ekstraksi fitur, yaitu kurtosis, mean, median, gradien, dan deviasi standar. Mean-med-grad ditentukan sebagai kombinasi terbaik, berdasarkan uji signifikansi ($p= 1{,}54 \times 10^{-6}$) dan prinsip parsimoni, dengan akurasi pada data uji sebesar 0,85 menggunakan model LDA. Sementara, akurasi tanpa seleksi fitur (mean-med-grad-std-kurt) sebesar 0,84. Validasi menggunakan GC-MS memperkuat hasil *e-nose*, di mana konfigurasi pemisahan kelas pada hasil analisa *e-nose* cukup bersesuaian dengan hasil analisa GC-MS dalam membedakan teh fermentasi dan non-fermentasi. Analisis GC-MS menunjukkan bahwa *methyl alcohol* dan *dimethyl sulfide* terdeteksi dalam kadar lebih tinggi pada teh non-fermentasi dibandingkan teh fermentasi, sementara *(1R)-4,7,7-trimethylbicyclo[2,2,1]heptan-2-one* hanya ditemukan pada teh fermentasi. GC-MS memberikan pembuktian kimia, sedangkan kombinasi fitur terpilih memastikan bahwa perbedaan kimia tersebut tercermin secara konsisten dalam representasi sinyal sensor. Hasil ini menunjukkan efektivitas seleksi fitur berbasis skor *silhouette* dalam meningkatkan performa klasifikasi *e-nose* dalam membedakan teh fermentasi dan non-fermentasi.

Kata kunci: *e-nose*, fermentasi, seleksi fitur, skor *silhouette*, teh.

ABSTRACT

Differences in fermentation degrees result in variations in tea quality characteristics, including changes in bioactive compounds, volatile compounds, taste, and aroma in green tea (non-fermented), yellow tea (lightly fermented), red tea (semi-fermented), and black tea (fully fermented). Therefore, accurate classification is essential to support quality control and enhance product competitiveness. Aroma, which reflects the profile of volatile compounds, is the focus of this study and is analyzed using an electronic nose (e-nose) system. In e-nose systems, the generated data are high-dimensional, making the selection of appropriate feature combinations crucial for improving classification performance. This study aims to develop a tea classification method based on e-nose data by optimizing feature selection using silhouette scores in the Linear Discriminant Analysis (LDA) space. E-nose data were analyzed using five feature extraction methods, namely kurtosis, mean, median, gradient, and standard deviation. The mean–median–gradient combination was identified as the optimal feature set based on significance testing ($p = 1.54 \times 10^{-6}$) and the principle of parsimony, achieving a test accuracy of 0.85 using the LDA model, compared to 0.84 without feature selection (mean-med-grad-std-kurt). Validation using GC-MS strengthened the e-nose results, as the class separation configuration obtained from e-nose analysis showed good agreement with GC-MS analysis in distinguishing fermented and non-fermented teas. GC-MS analysis revealed that methyl alcohol and dimethyl sulfide were detected at higher levels in non-fermented tea compared to fermented tea, whereas (1R)-4,7,7-trimethylbicyclo[2,2,1]heptan-2-one was found exclusively in fermented tea. GC-MS provides chemical evidence, while the selected feature combination ensures that these chemical differences are consistently reflected in the sensor signal representation. These results demonstrate the effectiveness of silhouette score-based feature selection in enhancing the classification performance of e-nose systems for distinguishing fermented and non-fermented teas.

Keywords: e-nose, fermentation, feature selection, silhouette score, tea.