Review of: "Infrared Spectroscopy (FT-NIR) and t-Distributed Stochastic Neighbor Embedding (t-SNE) as an Analytical Methodology for Rapid Identification of Tea Adulteration"

Tatas Hardo Panintingjati Brotosudarmo[1]

1 Universitas Ciputra Surabaya, Surabaya, Indonesia

Potential competing interests: No potential competing interests to declare.

Congratulations on developing a non-destructive method that utilizes FT-NIR and t-SNE for the rapid identification of tea adulteration. This approach holds great promise for practical application in the future. Below are my suggestions and comments.

Introduction part: Although the introduction to tea beverages is comprehensive, the presentation of the research questions and challenges could be enhanced by making the following adjustments:

1. Clearly articulate the research questions and frame them as compelling challenges or critical inquiries that warrant attention.
2. Emphasize the significance of the study by addressing why FT-IR and t-SNE are essential for advancing rapid analytical methods.
3. Clearly outline the evaluation criteria and considerations for testing to provide better clarity.

Methods part: More additional detail regarding the data collection and data processing needs to be improved. Information on the experimental design and the exact conditions of treatment during preparation needs to be described, including when and how the samples were collected, parameters for dehydrated samples, and how the samples were cut. Related to data handling, whenever feasible, sharing the data and/or code can significantly benefit the reproducibility of experiments conducted by other researchers. As this paper focuses on methodology, it would be beneficial to include a chart that visually represents the entire process.

Results and Discussion: In Fig 1b, 1c, and 1d, please ensure a consistent sequence for the order of samples: Standard Pieces, R1, R2, and R3, as well as maintaining a uniform color scheme throughout. For Fig 2, adding an arrow or another marker to highlight the band at 10,340.70 cm^{-1} would help guide the reader's attention and facilitate evaluation of the distinctions present. Additionally, it is important to discuss the parameters and rationale behind the deconvolution process to achieve an accurate fitting. In Figures 2a and 2d, as well as Figures 3a, 3b, and 3c, the filled areas representing the overall spectrum appear to lack a consistent baseline; please provide clarification on this aspect. In part 3.3, there is a statement, "t-SNE proved to be the most efficient in forming and identifying groups, possibly due to its non-linear clustering approach that considers spectral features not detected by PCA." This statement needs to be clarified and discussed in more detail.

I wish you all success for the publications!