IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Instrumentation Mesure Métrologie
Vol. 23, No. 6, December, 2024, pp. 441-450
Journal homepage: http://iieta.org/journals/i2m

Enhancing Ship Coating Quality Detection via Machine Learning-Optimized Visible Near-Infrared Spectroscopy

Ali Khumaidi[1], Ridwan Raafi'udin[2*], Nusa Setiani Triastuti[3]

[1] Department of Informatics, Krisnadwipayana University, Jakarta 13077, Indonesia
[2] Department of Informatics, Universitas Pembangunan Nasional Veteran Jakarta, Jakarta 12450, Indonesia
[3] Department of Civil Engineering, Krisnadwipayana University, Jakarta 13077, Indonesia

Corresponding Author Email: raafiudin@upnvj.ac.id

https://doi.org/10.18280/i2m.230604

Received: 1 October 2024
Revised: 2 December 2024
Accepted: 16 December 2024
Available online: 24 December 2024

Keywords:
classification, feature selection, non-destructive testing, portable spectroscopy, ship coating, spectral transformation, Vis/NIR

ABSTRACT

This study proposes a portable Visible/Near-Infrared (Vis/NIR) spectroscopy-based approach to detect and evaluate the quality of ship coatings. Vis/NIR spectroscopy offers an accurate, non-destructive method for identifying coating conditions through spectral data acquisition, combined with machine learning analysis to improve detection performance. In this study, using a device with a wavelength of 410-940 nm, spectral transformations such as scatter correction, baseline correction, smoothing, and derivative were applied to improve data quality, followed by feature selection using PCA and IFS. SVM, Random Forest (RF), and LDA classification algorithms were then used to model spectral data. The coating quality consists of four classes, with 40 samples for each. The initial results of modeling without treatment were improved with an average accuracy of 83.90%. Then, applying the combination of Nippy and IFS significantly increases average accuracy results by 96.86%. Incorporating spectral transformation and feature selection methods can optimally utilize spectral information and improve the model's overall performance with an increase in accuracy of 12.96%.

1. INTRODUCTION

Coating on ships is essential in the shipping industry to protect ship structures from corrosion, abrasion, and the growth of marine organisms [1]. Good quality coating can extend the ship's life, reduce maintenance costs, and improve operational efficiency. Given the harsh marine environment, regular inspection and monitoring of coatings is necessary to maintain the integrity of the vessel and prevent further damage [2]. Suppose coating quality is not taken care of. In that case, the impact can be devastating, including increased risk of corrosion, structural damage, and the growth of marine organisms, leading to increased operational costs [3]. This highlights the importance of standards and regulations such as those set by the International Maritime Organization (IMO) and the Indonesian Classification Bureau (BKI), which require regular inspection of coatings as part of the ship's certification and maintenance process [4]. Coating failure can cause serious structural damage and decrease the ship's performance [5]. Therefore, it is necessary to develop more sophisticated and accurate coating detection methods to ensure effective protection and maintain ships' safety and operational efficiency [6].

The commonly used methods of ship coating inspection so far include visual inspection, thickness measurement using a coating thickness gauge, and adhesion test to assess the adhesion of the coating on the surface of the ship [7]. Visual inspections are carried out to detect visible damage such as cracks, peeling, or rust, while thickness measurements ensure the coating has an adequate thickness as per the standard [8]. However, this method has limitations, such as the inability to detect micro-damage or delamination under the coating surface and reliance on the inspector's skill. In addition, these methods tend to be time-consuming and prone to human error, especially in subjective visual inspections [9]. Therefore, new technologies are needed that can overcome these constraints and offer a faster, more accurate, and non-destructive approach.

Visible/Near-Infrared (Vis/NIR) spectroscopy has developed as a promising alternative to non-destructive analysis. This technology can detect the chemical and physical properties of the coating layer without damaging it, allowing for the detection of micro-damage, delamination, or structural changes that conventional methods may not detect [10, 11]. In addition, Vis/NIR spectroscopy allows for fast measurements and can be applied in hard-to-reach areas, making it a more practical solution than conventional methods [12]. This approach has also been widely applied in various fields, such as materials analysis [13, 14], food [15], and pharmaceuticals [16]. However, the application of Vis/NIR spectroscopy in ship coating analysis is still relatively rarely studied, especially for micro-damage detection or delamination. When combined with machine learning analysis, this approach can improve accuracy and consistency in assessing coating quality, making it an effective solution in the shipping industry [17].

Despite its potential, Vis/NIR data processing often faces challenges such as noise, scatter, and baseline shift, which can affect the accuracy of predictive models [18]. Spectral transformation is an important stage in NIR processing because it can improve model performance [19]. The use of the best spectral transformation method is often determined through trial and error. Spectral transformations aim to eliminate all sources of informative variance from the spectral [20]. Several studies compare different spectral transformations to produce optimal model inputs [21]. Spectral transformations, such as clipping, scatter correction, smoothing, and derivatives, have been shown to improve data quality by eliminating distractions [22-24]. In addition to spectral transformation, feature selection is a crucial step in Vis/NIR data modeling and is able to improve relevant information signals [25]. Principal Component Analysis (PCA) plays a role in reducing the dimensions of data without sacrificing important information [26]. Iterative elimination feature selection has a fairly effective performance in improving the performance of meat quality prediction models [27]. The combination of spectral transformation and feature selection can improve the performance of machine learning models in detecting, classifying, and predicting material quality based on Vis/NIR data, making it a relevant approach in accurately detecting coating quality.

Until now, there are still limitations in fast, accurate, and non-destructive coating inspection methods, especially in detecting micro-damage or delamination that is not detected by conventional methods. The novelty of this study lies in the combination of portable Vis/NIR spectroscopy with machine learning analysis equipped with spectral transformation and feature selection to improve prediction accuracy. This research aims to develop a method that is able to detect micro-damage and evaluate coating quality quickly, accurately, and non-destructively. The specific goal of this study is to achieve an increase in detection accuracy of up to 10-15% compared to the basic method of processing Vis/NIR data without spectral transformation and feature selection while maintaining time efficiency and practicality of application in the field. With this approach, this research not only offers more sophisticated solutions but also contributes to developing more reliable and efficient coating inspection methods to support ships' safety and operational efficiency.

2. DEVICE HARDWARE AND SOFTWARE DEVELOPMENT

The tool used to acquire data in this study is the result of assembling. The main sensor used is the Sparkfun AS7265x [28] with specifications as shown in Table 1, while the main board used is a single board computer from Raspberry, namely the Raspberry Pi Zero W with specifications shown in Table 2. The schematic design of the tool assembly can be seen in Figure 1. The AS7265x sensor is connected to a Raspberry Pi using four cables with the I2C communication protocol [29].

The number of pages for the manuscript must be no more than ten, including all the sections. Please make sure that the whole text ends on an even page. Please do not insert page numbers. Please do not use the Headers or the Footers because they are reserved for technical editing by editors.

I2C (Inter-Integrated Circuit) communication is a serial protocol widely used to connect microcontrollers with peripheral devices such as sensors. The Raspberry Pi is connected to the AS7265x sensor via I2C communication in the image. The red wire connects the 3.3V pin of the Raspberry Pi to the *VCC* pin of the AS7265x sensor, providing a 3.3V power supply, while the black wire connects the *GND* pins to ensure a common ground reference. The blue wire connects the *SCL* (Serial Clock) pins to transmit the clock signal, which synchronizes data transfer between the Raspberry Pi and the sensor. The green wire connects the *SDA* (Serial Data) pins for bidirectional data communication.

Table 1. AS7265X sensor specification

Specification	Description
Sensor Type	Multi-spectral sensor system
Spectral Bands	UV (410 nm, 435 nm), VIS (460 nm, 485 nm, 510 nm, 535 nm, 560 nm, 585 nm), NIR (610 nm, 645 nm, 680 nm, 705 nm, 730 nm, 760 nm, 810 nm, 860 nm, 900 nm, 940 nm)
Optical Interface	I2C and UART communication
Operating Voltage	3.3V (typical)

Table 2. Raspberry Pi Zero W specification

Specification	Description
Processor	BCM2835, ARM11 core, 1 GHz
RAM	512 MB LPDDR2 SDRAM
Wi-Fi	802.11 b/g/n
Operating System	Raspberry Pi OS Headless
Power Supply	5V via Micro-USB

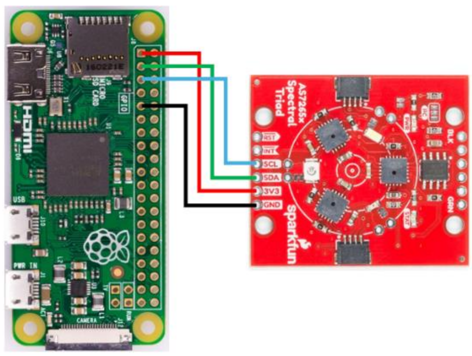


Figure 1. Schematic of the design of the tool

In this setup, the Raspberry Pi acts as the master, controlling the communication by generating the clock signal and sending or receiving data through the *SDA* line. The AS7265x sensor acts as the slave, responding to the master's requests. The communication process begins with the Raspberry Pi sending a *start condition* and the address of the AS7265x sensor to identify it on the I2C bus. Once the sensor recognizes its address, data is transferred via the *SDA* line in synchronization with the clock signal on the *SCL* line. The communication concludes with a *stop condition*. This setup demonstrates the advantages of I2C communication, including its simplicity in requiring only two data lines and its capability to connect multiple devices on the same bus, provided each device has a unique address. The configuration shown provides an efficient way to read data from the AS7265x sensor using the Raspberry Pi.

The calibration and validation of the spectral measurements were conducted using a Raspberry Pi Zero W microcontroller paired with the AS7265x spectral sensor, which operates via the I2C communication protocol. For calibration, a known light source with a stable and well-characterized spectral output was employed. The AS7265x sensor readings were compared to the reference values from a calibrated spectrometer to ensure accuracy. Adjustments were made by applying calibration coefficients to correct for sensor-specific deviations. For validation, the sensor was tested against various light sources, including natural sunlight, LED lights, and fluorescent bulbs, to verify the consistency and accuracy of the spectral data across a range of conditions. These procedures ensured reliable and reproducible spectral measurements for the intended applications.

The data generated by the sensor is in the form of spectroscopy with a wavelength of 18 pieces, as shown in Figure 2. The 18 data consists of three groups of reflectance types represented by their respective colors. Blue is for the ultraviolet (UV) light group, red is for the visible light group (VIS), and green is for the near-infrared (NIR) group.

The programming language used for data acquisition application development is Python with the AS7265x library from Github [30]. At the same time, the main programming language for modeling is Python, with some additional libraries such as Numpy, Pandas, and Matplotlib.

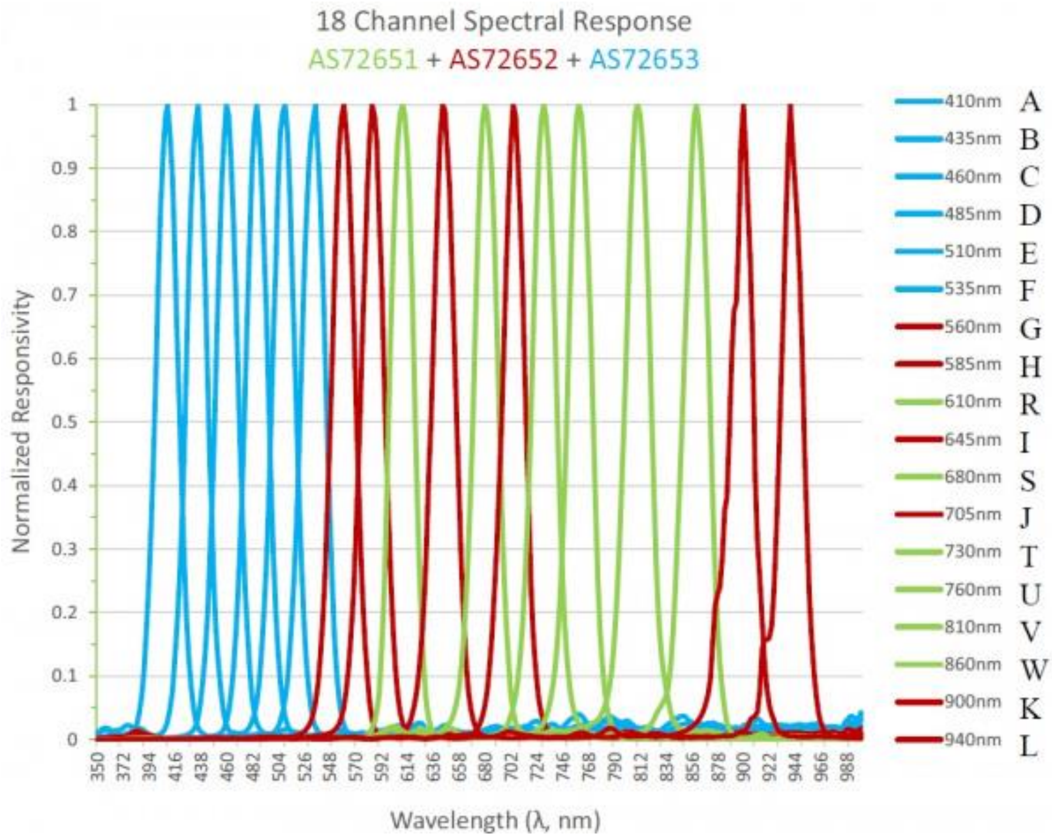


Figure 2. Wavelength range

3. MATERIALS AND METHODS

3.1 Ship coating sample

The assessment of the condition of the coating of the ship in the context of the Condition Assessment Program (CAP), which is regulated by the Indonesian Classification Bureau (BKI), includes several categories that describe the condition of the coating system on the structure of the ship. This category consists of four main classifications [31], namely:

(1) Very Good Condition, the coating system is in optimal condition and does not require maintenance or repair. The allowable corrosion margin is more than 50%.

(2) Good Condition: The coating system is in good condition with maintenance and documentation, without the need for repairs. The remaining corrosion margin is between 25% to 50%.

(3) Class Condition: the coating system is in poor condition, but the corrosion protection system still functions well. Maintenance and documentation are considered satisfactory, but the allowable margin of corrosion remains less than 25%.

(4) Poor condition: the coating system is in poor condition, and the corrosion protection system is also in poor condition.

Maintenance or repair is required to restore serviceability, as the allowable corrosion margin has been exceeded. In this study, as many as 40 coating samples were taken from each coating class representing the four categories. Sampling was conducted on Roll-on/Roll-off (RoRo) and Ferry ships, which were selected for their unique operational characteristics and exposure to marine environmental conditions. RoRo vessels often operate at high frequencies and face variations in mechanical loads that can affect coating conditions, especially in ramp areas and vehicle decks.

Meanwhile, ferries, which generally serve inter-island routes with tight schedules, are also exposed to high humidity, saltwater spray, and significant temperature variations. This makes both types of vessels representative in evaluating coating performance in challenging marine environments while providing relevant data to develop more reliable and accurate inspection methods. Prior to sampling, the areas that the experts have determined are then marked and cleaned to ensure accurate and representative results. The coating materials used include epoxy and polyurethane-based coatings, which are commonly used in the marine industry.

3.2 Spectral acquisition

Portable Vis/NIR spectroscopy devices are pre-calibrated before being used to collect reflectance spectral data from the ship's marked coating points. The integration time was set at 100 ms, and each sample was measured three times, with the average of the three measurements used for further analysis.

Each measurement is performed under controlled lighting conditions in a shaded area to minimize the effects of ambient light and improve the reliability of spectral data. The device is placed perpendicular to the coating surface to avoid measurement bias due to angular reflectance variations. Environmental conditions during data collection, such as temperature (25-30°C) and relative humidity (60-75%), are monitored and maintained within a consistent range. The process of acquiring this data is illustrated in Figure 3.

3.3 Proposed methods

In this study, an approach that combines spectral transformation and feature selection before modeling is proposed to improve the accuracy of the classification of ship coating conditions. Spectral transformation is applied as a first step to improve data quality by reducing noise, eliminating baseline effects, and amplifying relevant spectral information. Feature selection is carried out to identify the most significant wavelengths in separating the different coating classes. By combining spectral transformation and feature selection, it is hoped that the classification model built can work optimally in detecting and classifying coating conditions more accurately. The process of developing a classification model can be seen in Figure 4.

3.3.1 Spectral transformation

Spectral transformations are applied to improve spectral data quality before being incorporated into the classification model. In this study, four spectral transformation methods were used, namely:

(1) scatter correction, which functions to correct light scattering that may appear during measurement;

(2) normalization aimed at equalizing spectral data into a uniform range;

(3) baseline correction to eliminate unwanted baseline effects from spectral data;

(4) smoothing and derivation using the Savitzky-Golay (SAVGOL) method. This SAVGOL uses several parameters, namely 'filter_win', 'poly_order', and 'deriv_order'.

This transformation aims to reduce noise in spectral data and extract more subtle spectral features that are relevant for modeling. More details on the transformations used can be seen in Table .

Figure 3. Spectral data acquisition

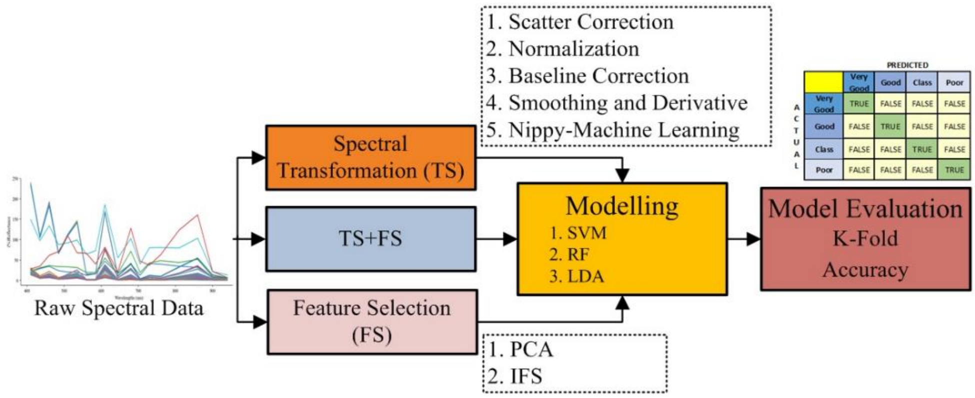


Figure 4. Proposed coating classification model

Table 3. Methods, operations, parameters, and values of spectral transformations

Method/Operator	Parameter	Values	Impact
SNV	iqr	75-25, 90-10	Removes light scattering effects by normalizing the spectral with its standard deviation.
RNV			Corrects scattering using the interquartile range, reducing outliers appearing at spectral edges.
MSC			Adjusts spectra by reducing baseline effects and linear scaling to correct scattering variations.
NORML			Standardizes spectral intensity within a uniform range, improving consistency across samples.
BASELINE			Removes unwanted baseline components from the spectral, enhancing the detection of small but relevant features.
SAVGOL	filter_win	5, 7, 11	Controls the filter window size to suppress noise; larger windows smooth more but may reduce spectral detail.
	poly_order	3	Sets the polynomial degree for fitting, enhancing flexibility in capturing spectral patterns.
	deriv_order	1, 2	The first derivative highlights intensity changes, while the second derivative emphasizes inflection points for specific features.

3.3.2 Feature selection

Feature selection is applied to reduce the spectral data dimensions and improve the performance of the classification model. The two feature selection methods used are principal component analysis (PCA) and iterative elimination feature selection (IFS). PCA is a dimensionality reduction technique that projects data into key components that maximize variation, resulting in a more concise representation of data without significant information loss. This method is used to capture important patterns in the spectra that can affect model performance, especially in datasets with multicorrelation between wavelengths [32].

IFS is based on line simplification with an angular elimination system [27]. The collected spectroscopy data is converted into a line shape, and then the line is simplified by removing the corners from the line. This process starts from the smallest corner to the largest corner. Each iteration of the simplification process will eliminate one corner, reducing one column of data in the corresponding dataset. This approach aims to reduce the risk of overfitting, speed up the training process, and improve the model's generalization on the test data.

3.3.3 Modelling

In this study, several classification algorithms are applied to model spectral data and predict the quality of ship coatings, namely:

(1) Support Vector Machine (SVM) with linear kernel. The linear kernel on the SVM allows the modeling of the linear relationships between features in spectral data [33].

(2) Random Forest Classifier (RF) is a decision tree-based ensemble method. The RF algorithm randomly selects a subset of features and data samples at each iteration of tree formation, aiming to improve resistance to overfitting and improve prediction accuracy [34].

(3) Linear Discriminant Analysis (LDA) works by looking for linear projections of features that maximize separation between classes.

This algorithm maximizes the ratio between variation between classes and variation within classes in the projected feature space. These three algorithms are implemented by dividing training and testing data with a ratio of 80:20 using 5-fold cross-validation to evaluate the model's performance.

3.3.4 Model evaluation

Classification accuracy describes the percentage of correctly classified samples of the overall sample and is used as the primary metric to measure model performance [35]. However, when the distribution of the number of samples between classes is unbalanced, the confusion matrix provides a more comprehensive picture of the performance of the classification model by providing information about the correct classification and the model's predictions for a particular category [36]. For example, if the poor class is identified as positive, then the very good, good, and class classes are set as negative. In this context, True Positive (TP) refers to samples that are indeed positive and correctly classified as poor. True Negative (TN) is a negative sample that is correctly classified as a class other than poor. False Positives (FP) occur when a negative sample is incorrectly classified as poor, while False Negative (FN) occurs when a positive sample is incorrectly placed into another negative category, for example, when a sample that should be classified as good is incorrectly classified as very good or class. The performance evaluation of this classification model mainly uses classification accuracy measured based on the confusion matrix, in the case of ship coating classification.

4. RESULTS AND DISCUSSION

4.1 Spectral acquisition results

Spectral data collection in the field produced spectral data of 40 very good classes, 33 good classes, 37 class classes, and 38 poor classes, as shown in Figure 5 and Table 4. The average spectral of the ship's coating data without spectral transformation is illustrated in Figure 5. Spectroscopy Data Plotting has 18 different wavelength points in the wavelength range of 410–940 nm with several peaks and valleys of the wave. When NIR radiation hits a sample, the phenomenon of absorption, reflection, and transmission occurs, where the phenomenon depends on the chemical elements that make up the sample [37]. The visualization results of the spectral data that were acquired late can be seen in Figure 5.

Table 4. Data distribution

Coating Class	Amount of Data
Very Good	40
Good	33
Class	37
Poor	38
Total	148

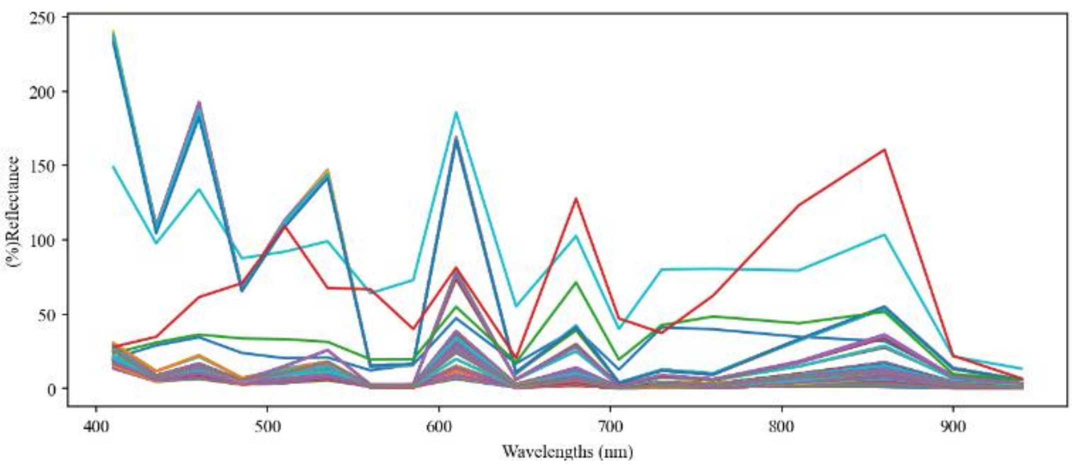


Figure 5. Spectroscopy data plotting

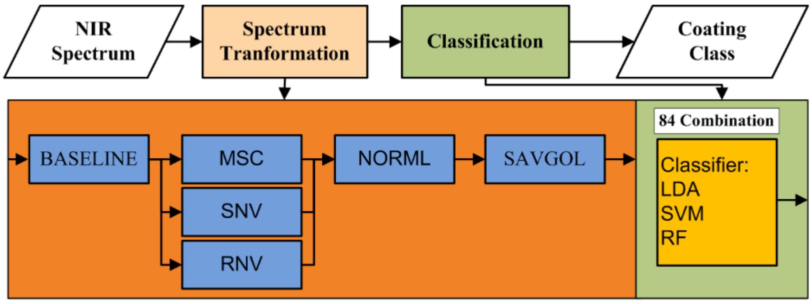


Figure 6. Flow determination of the best spectral transformation according to modeling

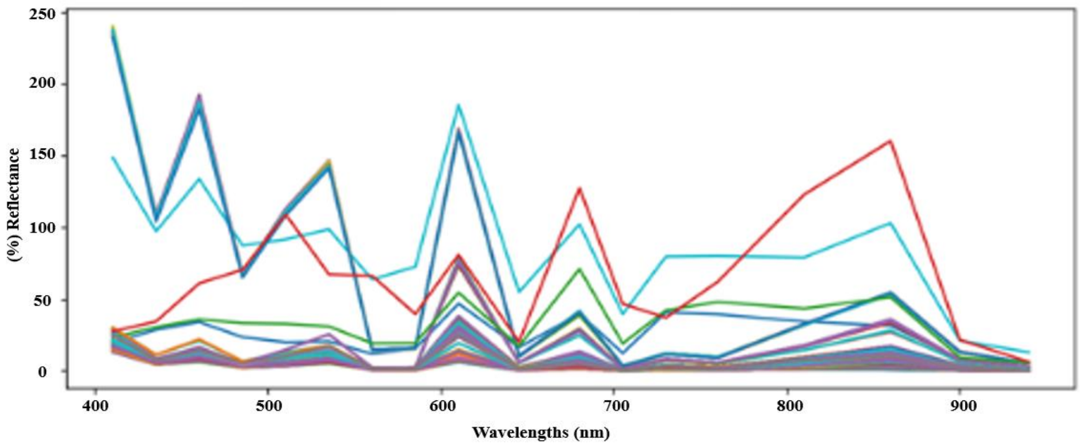
4.2 Spectral transformation

Machine learning was used to determine the best spectral transformation strategy by testing combinations of 6 transformation operators and related parameters (Table 3), resulting in a total of 84 combinations. This process involves automated learning and systematic tuning of hyperparameters. The collection and evaluation of spectral transformation operator combinations are carried out based on the performance of the generated model, where the comparison of the effects of different operator combinations is assessed from the model performance (Figure 6). This combination of spectral transformations is evaluated using the highest accuracy on classification algorithms such as LDA, SVM, and RF for the classification approach.

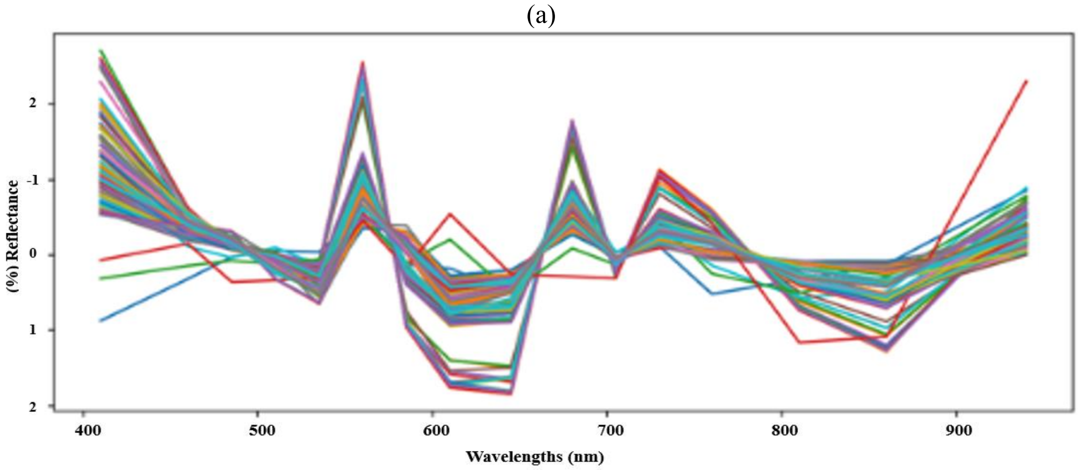
Table 5 shows a comparison of the performance of several spectral transformation methods using three classification algorithms: SVM, RF, and LDA. From these results, the Nippy spectral transformation method results in the highest performance for all algorithms, with the highest accuracy values on SVM (94.55%), RF (96.64%), and LDA (97.33%). In contrast, the approach without spectral transformation (None) showed lower performance, especially in SVM (70.13%). The use of the First and Second Derivative methods also gives lower yields on SVMs but remains strong on RF and LDA. Other spectral transformations such as MSC, SNV, Baseline Correction, and SAVGOL provide a variation in results, with SAVGOL and MSC showing competitive performance on RF and LDA.

Table 5. Spectral transformation performance comparison

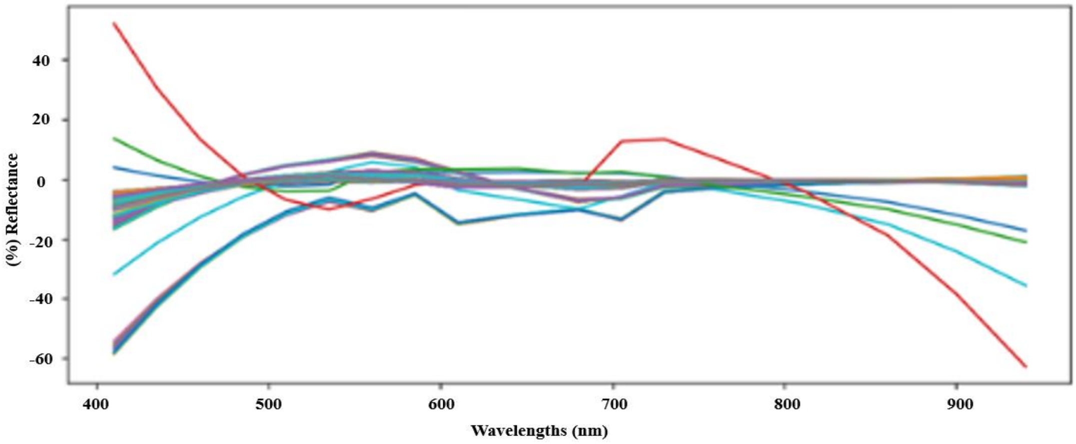
Spectral Transformation	SVM	RF	LDA
None	70.13%	89.10%	92.48%
MSC (multiplicative scatter correction)	76.96%	84.45%	92.50%
SNV (standard normal variate)	74.96%	81.65%	91.21%
First dan Second Derivative	49.24%	92.52%	92.55%
Baseline Correction	68.29%	89.10%	92.48%
SAVGOL ('filter_win': 5, 'poly_order': 3)	65.33%	93.17%	92.48%
Nippy	94.55%	96.64%	97.33%



(a)



(b)



(c)

Figure 7. Spectral after transformation, (a) SVM, (b) RF, (c) LDA

The results of data processing with spectral transformation using the Nippy module show that the combination of different transformation operators can significantly improve the accuracy of the classification model. In the SVM algorithm, the best operator combination is SAVGOL with parameters {'deriv_order': 1, 'filter_win': 5, 'poly_order': 3} and BASELINE. In the LDA algorithm, the best operator combination is SAVGOL with parameters {'deriv_order': 1, 'filter_win': 11, 'poly_order': 3}, BASELINE and NORML. Meanwhile, for the RF algorithm, the combination of RNV operators with parameters {'iqr': [75.0, 25.0]} and SAVGOL with parameters {'deriv_order': 2, 'filter_win': 5, 'poly_order': 3. An illustration of the spectral after transformation can be seen in Figure 7.

4.3 Feature selection and combined methods

Table 6 shows the performance comparison between the use of feature selection and the combination of feature selection with spectral transformation on SVM, RF, and LDA algorithms. Across all algorithms, the application of IFS showed a significant performance improvement over PCA, with the highest accuracy in SVM (96.62%), RF (91.10%), and LDA (94.57%). The combination of spectral transformations using the Nippy module with a selection of features further improves the model's performance. The combination of Nippy and IFS provides the best results, especially on RF (96.64%) and LDA (97.33%) algorithms, and maintains optimal performance on SVM (96.62%).

Table 6. Comparison of feature selection performance and spectral transformation

Methods	SVM	RF	LDA
PCA	92.48%	90.43%	89.24%
IFS	96.62%	91.10%	94.57%
Nippy+PCA	94.55%	93.86%	95.26%
Nippy+IFS	96.62%	96.64%	97.33%

Table 7 shows the standard deviation values of three different methods: SVM, RF, and LDA. The standard deviation value is obtained from the results of combining Table 5 and Table 6. The smaller standard deviation values indicate that the method has a smaller and more consistent variation in results. Therefore, the LDA method with the smallest standard deviation shows the most consistent results compared to SVM and RF. It is important to consider in the selection of the right method according to the needs of the analysis.

Table 7. Standard deviation comparison of algorithms

Algorithms	SVM	RF	LDA
Standard Deviation	0.153	0.044	0.024

The bar graph in Figure 8 shows that the Nippy and IFS methods significantly improve the mean accuracy compared to the None method. The combination of the two methods (Nippy+IFS) produced the highest mean accuracy of 96.86%. This shows that the merger of methods can have a synergistic impact in improving classification accuracy. Individually, the Nippy method recorded the most significant improvement compared to IFS, with a mean accuracy of 96.17% and 94.10%, respectively.

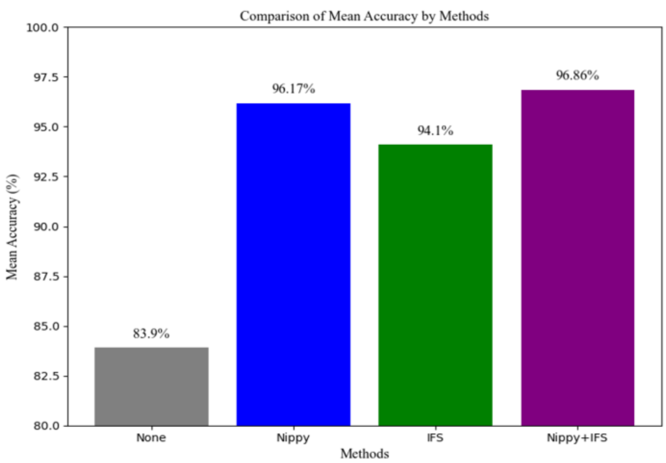


Figure 8. Comparison of mean accuracy by methods

This indicates that the Nippy approach-based method has an advantage in perfecting the classification results. However, even though IFS provides lower performance than Nippy, the combination of the two has proven to produce the highest accuracy.

These findings emphasize the importance of the exploration and development of a combined approach in classification algorithms. By leveraging the strengths of each method, Nippy+IFS is able to deliver superior results, demonstrating the potential for synergy in the combination of complementary techniques.

4.4 Discussion

Spectral transformation has proven to play an important role in improving the accuracy of classification models. Nipp's method that applies machine learning to find the best operator quickly and consistently gives the best results among all algorithms, according to research by Khumaidi et al. [17, 38], which shows that spectral transformation is able to optimize spectral information better than other methods. In contrast, the no-transform approach gives lower yields, especially on SVMs, which shows that the model cannot effectively capture data variations without spectral adjustments. Interestingly, the first and second derivative transformations show a significant performance degradation in SVMs, which may be due to overfitting to more subtle spectral changes. This indicates that the selection of the right transformation is highly dependent on the characteristics of the data and the classification algorithm used. The spectral transformation operators that improve performance on all three classifiers are SAVGOL and BASELINE.

Feature selection is crucial in improving classification accuracy, especially when combined with spectral transformation. The use of IFS consistently performs better than PCA, suggesting that iterative feature selection elimination is more effective in identifying the most relevant features [27]. In addition, the combination of spectral transformations with Nippy shows significant performance improvements across all algorithms, especially when combined with IFS, which provides the highest accuracy on RF and LDA. This shows that incorporating spectral transformation and feature selection methods can optimally utilize spectral information and improve the model's overall performance.

To improve understanding of the limitations and applications of the technology in real-world scenarios, it is important to conduct thorough testing and gather feedback from practical implementations. It is also important to consider several factors that can affect the performance of spectral transformation in a practical context. One of the main limitations is the complexity of selecting the right transformation operator, which is highly dependent on the specific characteristics of the data and the classification algorithm used. For example, although Nippy was shown to deliver the best consistent results in this study, in real-world applications, data processing at large scales or varying environmental conditions can introduce challenges in selecting the optimal transformation. In addition, spectral transformations such as first and second derivatives that cause performance degradation in SVMs may reveal overfitting problems in real data with wider variation or higher noise, which is not necessarily resolved by the tested method.

Pointing to future research, these findings open up several opportunities for further exploration, particularly in developing more robust techniques to address these challenges and improve practical outcomes. Further research can be focused on developing classification algorithms that are more resistant to overfitting on subtle spectral variations, as well as feature selection methods that are more adaptive to changes in data conditions. Exploring the use of deep learning techniques or hybrid approaches between spectral transformations and neural network-based models can provide a deeper understanding of how to optimize spectral data processing in real-world applications. In addition, future research may examine the application of spectral transformation in a broader domain, such as environmental monitoring or image-based disease detection, to test the sustainability and generalization of the findings.

5. CONCLUSION

This study shows that the combination of spectral transformation methods and machine learning-based feature selection has great potential in improving the accuracy of ship coating quality detection using portable VIS/NIR spectroscopy. The proposed approach not only addresses the challenges in spectral data processing, such as noise and signal variation but also optimizes the performance of the classification model. The results of the comparison between SVM, Random Forest, and LDA algorithms indicate that the right strategy in spectral processing and analysis is essential to achieve reliable prediction results. This method offers an effective solution for applications in the shipping industry, opening up further opportunities to develop fast, accurate and non-destructive inspection systems. Applying this solution in the shipping industry can result in significant efficiencies in terms of time and cost. Faster and more accurate inspection of coating quality will minimize ship downtime, improve safety, and extend the vessel's service life. In addition, this approach has great commercial potential, allowing for the development of portable inspection tools that can be used extensively by various parties in the shipping industry. This can create new business opportunities in the field of developing more advanced and reliable ship inspection and maintenance technology.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the Directorate of Research, Technology, and Community Service (DRTPM), Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia, for providing financial support through research Grant Number: 105/E5/PG.02.00.PL/2024. Special thanks are also extended to PT Biro Klasifikasi Indonesia and PT Dok dan Perkapalan Kodja Bahari Galangan Jakarta II for their invaluable support and facilitation during the research process. The contributions and collaboration from all parties were instrumental to the success of this research.

REFERENCES

[1] Thomas, P., Sahoo, B.N., Thomas, P.J., Greve, M.M. (2024). Recent advances in emerging integrated anticorrosion and antifouling nanomaterial-based coating solutions. Environmental Science and Pollution Research, 1-27. https://doi.org/10.1007/s11356-024-33825-6

[2] Tadros, M., Ventura, M., Guedes Soares, C. (2023). Review of the decision support methods used in optimizing ship hulls towards improving energy efficiency. Journal of Marine Science and Engineering, 11(4): 835. https://doi.org/10.3390/jmse11040835

[3] Vedaprakash, L., Senthilkumar, P., Inbakandan, D., Venkatesan, R. (2022). Marine biofouling and corrosion on long-term behavior of marine structures. In A Treatise

on Corrosion Science, Engineering and Technology. Singapore: Springer Nature Singapore, pp. 447-466. https://doi.org/10.1007/978-981-16-9302-1_24
[4] Momber, A.W., Langenkämper, D., Möller, T., Nattkemper, T.W. (2023). The exploration and annotation of large amounts of visual inspection data for protective coating systems on stationary marine steel structures. Ocean Engineering, 278: 114337. https://doi.org/10.1016/j.oceaneng.2023.114337
[5] Momčilović, N., Ilić, N., Kalajdžić, M., Ivošević, Š., Petrović, A. (2024). Effect of corrosion-induced structural degradation on the ultimate strength of a high-tensile-steel ship hull. Journal of Marine Science and Engineering, 12(5): 745. https://doi.org/10.3390/jmse12050745
[6] Imran, M.M.H., Jamaludin, S., Ayob, A.F.M., et al. (2023). Application of artificial intelligence in marine corrosion prediction and detection. Journal of Marine Science and Engineering, 11(2): 256. https://doi.org/10.3390/jmse11020256
[7] Stojanović, I., Farkas, A., Alar, V., Degiuli, N. (2019). Evaluation of the corrosion protection of two underwater coating systems in a simulated marine environment. The Journal of The Minerals, Metals & Materials Society, 71(12): 4330-4338. https://doi.org/10.1007/s11837-019-03669-4
[8] Muzika, L., Svantner, M., Honner, M., Houdkova, S. (2021). Laser scanning thermography for coating thickness inspection. Engineering Proceedings, 8(1): 17. https://doi.org/10.3390/engproc2021008017
[9] Hütten, N., Alves Gomes, M., Hölken, F., Andricevic, K., Meyes, R., Meisen, T. (2024). Deep learning for automated visual inspection in manufacturing and maintenance: A survey of open-access papers. Applied System Innovation, 7(1): 11. https://doi.org/10.3390/asi7010011
[10] Luo, Y., Liu, Y., Wei, Q., Strlič, M. (2024). NIR spectroscopy in conjunction with multivariate analysis for non-destructive characterization of Xuan paper. Heritage Science, 12(1): 175. https://doi.org/10.1186/s40494-024-01287-1
[11] Li, X., Zhang, L.X., Zhang, Y., Wang, D., Wang, X.F., Yu, L., Zhang, W., Li, P.W. (2020). Review of NIR spectroscopy methods for nondestructive quality analysis of oilseeds and edible oils. Trends in Food Science & Technology, 101: 172-181. https://doi.org/10.1016/j.tifs.2020.05.002
[12] Sohn, S.I., Pandian, S., Oh, Y.J., Zaukuu, J.L.Z., Kang, H.J., Ryu, T.H., Cho, W.K., Cho, Y.S., Shin, E.K., Cho, B.K. (2021). An overview of near infrared spectroscopy and its applications in the detection of genetically modified organisms. International Journal of Molecular Sciences, 22(18): 9940. https://doi.org/10.3390/ijms22189940
[13] Liu, K., Fan, P.P., Jia, Z.C., Wang, Z.J., Qi, S.P. (2024). Analysis of four heavy metal concentrations in sediments from the Jiaozhou Bay, China by visible and near infrared spectroscopy (225–975 nm). Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 316: 124367. https://doi.org/10.1016/j.saa.2024.124367
[14] Zhang, K., Kwadzokpui, B.A., Adade, S.Y.S.S., Lin, H., Chen, Q. (2024). Quantitative and qualitative detection of target heavy metals using anti-interference colorimetric sensor Array combined with near-infrared spectroscopy. Food Chemistry, 459: 140305. https://doi.org/10.1016/j.foodchem.2024.140305
[15] León-Ecay, S., López-Campos, Ó., López-Maestresalas, A., Insausti, K., Schmidt, B., Prieto, N. (2024). Using portable visible and near-infrared spectroscopy to authenticate beef from grass, barley, and corn-fed cattle. Food Research International, 198: 115327. https://doi.org/10.1016/j.foodres.2024.115327
[16] Wu, X.H., Wang, Y.X., Wu, B., Sun, J. (2024). Classification of Fritillaria using a portable near-infrared spectrometer and fuzzy generalized singular value decomposition. Industrial Crops and Products, 218: 119032. https://doi.org/10.1016/j.indcrop.2024.119032
[17] Khumaidi, A., Purwanto, Y.A., Sukoco, H., Wijaya, S.H. (2022). Using fuzzy logic to increase accuracy in mango maturity index classification: Approach for developing a portable near-infrared spectroscopy device. Sensors, 22(24): 9704. https://doi.org/10.3390/s22249704
[18] Gullifa, G., Barone, L., Papa, E., Giuffrida, A., Materazzi, S., Risoluti, R. (2023). Portable NIR spectroscopy: The route to green analytical chemistry. Frontiers in Chemistry, 11: 1214825. https://doi.org/10.3389/fchem.2023.1214825
[19] Wang, Y.X., Yang, S., Yan, X.B., et al. (2023). Evaluation of data pre-processing and regression models for precise estimation of soil organic carbon using Vis–NIR spectroscopy. Journal of Soils and Sediments, 23(2): 634-645. https://doi.org/10.1007/s11368-022-03337-2
[20] Torniainen, J., Afara, I.O., Prakash, M., Sarin, J.K., Stenroth, L., Töyräs, J. (2020). Open-source python module for automated preprocessing of near infrared spectroscopic data. Analytica Chimica Acta, 1108: 1-9. https://doi.org/10.1016/j.aca.2020.02.030
[21] Xu, L., Zhou, Y.P., Tang, L.J., Wu, H.L., Jiang, J.H., Shen, G.L., Yu, R.Q. (2008). Ensemble preprocessing of near-infrared (NIR) spectra for multivariate calibration. Analytica Chimica Acta, 616(2): 138-143. https://doi.org/10.1016/j.aca.2008.04.031
[22] Lee, L.C., Liong, C.Y., Osman, K., Jemain, A.A. (2016). Effects of scatter-correction pre-processing methods and spectral derivative algorithms on forensic classification of paper. In Advances in Industrial and Applied Mathematics: Proceedings of 23rd Malaysian National Symposium of Mathematical Sciences (SKSM23), Johor Bahru, Malaysia. https://doi.org/10.1063/1.4954618
[23] Wu, M.J., Huang, Y.Q., Zhao, X., Jin, J., Ruan, Y.C. (2024). Effects of different spectral processing methods on soil organic matter prediction based on VNIR-SWIR spectroscopy in karst areas, Southwest China. Journal of Soils and Sediments, 24(2): 914-927. https://doi.org/10.1007/s11368-023-03691-9
[24] Gautam, R., Vanga, S., Ariese, F., Umapathy, S. (2015). Review of multidimensional data processing approaches for Raman and infrared spectroscopy. EPJ Techniques and Instrumentation, 2: 1-38. https://doi.org/10.1140/epjti/s40485-015-0018-6
[25] Zhang, Z.Y., Li, Y.X., Li, C.X., Wang, Z.C., Chen, Y. (2022). Algorithm of stability-analysis-based feature selection for NIR calibration transfer. Sensors, 22(4): 1659. https://doi.org/10.3390/s22041659
[26] Liu, X. (2020). Discriminative principal component analysis for high dimensional classification with applications in NIR spectroscopy. Doctoral dissertation, University College London.

[27] Raafi'udin, R., Purwanto, Y.A., Sitanggang, I.S., Astuti, D.A. (2024). Feature selection model development on near-infrared spectroscopy data. International Journal of Advanced Computer Science and Applications, 15(1): 645-653. https://doi.org/10.14569/IJACSA.2024.0150163
[28] SparkFun. (2018). AS7265x Datasheet. AMS Datasheet, 1(4): 1-63. https://cdn.sparkfun.com/assets/c/2/9/0/a/AS7265x_Datasheet.pdf.
[29] Mankar, J., Darode, C., Trivedi, K., Kanoje, M., Shahare, P. (2014). Review of I2C protocol. International Journal of Research in Advent Technology, 2(1): 474-479.
[30] GitHub. (2024). makoike/SparkFun_AS7265x_Python_Lib. SparkFun Triad Spectroscopy Sensor-AS7265x (Qwiic). https://github.com/makoike/SparkFun_AS7265x_Python_Lib.
[31] Biro Klasifikasi Indonesia. (2015). Part 1: Seagoing Ships. In Guidelines for Condition Assessment Program. Vol. 11, 2015 Ed.
[32] Ren, S., Jia, Y. (2023). Near-infrared data classification at phone terminal based on the combination of PCA and CS-RBFSVC algorithms. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 287: 122080. https://doi.org/10.1016/j.saa.2022.122080
[33] Huo, X.S., Chen, P., Li, J.Y., Xu, Y.P., Liu, D., Chu, X.L. (2024). Comparative study of linear and nonlinear calibration algorithm for extrapolation problem of near infrared spectroscopy quantitative analysis. Vibrational Spectroscopy, 132: 103693. https://doi.org/10.1016/j.vibspec.2024.103693
[34] Shahare, Y.R., Singh, M.P., Singh, S.P., Singh, P., Diwakar, M. (2024). ASUR: Agriculture soil fertility assessment using random forest classifier and regressor. Procedia Computer Science, 235: 1732-1741. https://doi.org/10.1016/j.procs.2024.04.164
[35] Qiu, Q., Nian, Y.J., Guo, Y., Tang, L., Lu, N., Wen, L.Z., Chen, D.F., Liu, K.J. (2019). Development and validation of three machine-learning models for predicting multiple organ failure in moderately severe and severe acute pancreatitis. BMC Gastroenterology, 19: 1-9. https://doi.org/10.1186/s12876-019-1016-y
[36] Chong, I.G., Jun, C.H. (2005). Performance of some variable selection methods when multicollinearity is present. Chemometrics and Intelligent Laboratory Systems, 78(1-2): 103-112. https://doi.org/10.1016/j.chemolab.2004.12.011
[37] Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biology and Technology, 46(2): 99-118. https://doi.org/10.1016/j.postharvbio.2007.06.024
[38] Khumaidi, A., Raafi'udin, R. (2022). Effects of oversampling smote and spectral transformations in the classification of mango cultivars using near-infrared spectroscopy. International Journal on Advanced Science Engineering Information Technology, 12(3): 1047-1053. https://doi.org/10.18517/ijaseit.12.3.16001