Petroleum Research 10 (2025) 622–635

Contents lists available at ScienceDirect

Petroleum Research

journal homepage: http://www.keaipublishing.com/en/journals/petroleum-research/

Full Length Article

A comparative experimental study on the prediction of renewable oils properties using RGB and HSV image processing techniques

Aditya Kolakoti [a, *], Ruthvik Chandramouli [b]

[a] School of Marine Engineering and Technology, Indian Maritime University, Kolkata Campus, West Bengal 700088, India
[b] Hindustan Aeronautics Limited, Avionics Division, Amethi, Uttar Pradesh, 227412, India

ARTICLE INFO

Article history:
Received 21 July 2024
Received in revised form
25 February 2025
Accepted 25 February 2025
Available online 4 March 2025

Keywords:
Red green blue (RGB)
Hue saturation value (HSV)
Image processing (IP)
Renewable oil properties

ABSTRACT

In this study, renewable oil properties of Flash Point (^{0}C), Fire Point (^{0}C), Density (kg/m^3), Cloud Point (^{0}C), Pour Point (^{0}C), and Viscosity (cST) are predicted using image processing techniques of Red Green Blue (RGB) and Hue Saturation Value (HSV). Eleven types of renewable oils are chosen for experimentation, and their surface images are captured with a high-resolution digital camera. For better accuracy, around 150 surface images are captured for each oil sample, and their average pixel data is extracted using RGB and HSV techniques. The digital pixel information (metadata) of all the oil samples is mapped to their experimental oil properties, and the accuracy of the developed metadata is validated with Fiji software due to its better image analysis and also complex data quantifying capabilities. The minimum, maximum, mean, mode and standard deviation results of RGB and HSV agree with Fiji. In addition, the developed dataset has been validated with Neural Network classification and TreeBagger algorithms. The results of TreeBagger reveal that the trained dataset is highly accurate (91.9% for RGB and 95.3% for HSV). Similarly, 95.6% (RGB) and 97.3% (HSV) accuracy is achieved for Neural Network classification. Finally, two new oil surface images are trained using the developed dataset. Both RGB and HSV accurately predict the oil properties. Therefore, it is evident that predicting the significant oil properties helps optimize the production process by reducing experimental costs and time.

1. Introduction

Renewable and sustainable energy sources are increasingly viewed as promising alternatives to conventional energy sources. They promote cleaner production and combustion, resulting in significantly less environmental pollution compared to conventional sources (Tamilselvan and Sudhakar, 2024). Owing to the substantial benefits of incorporating renewable energy sources across various applications, their demand has steadily risen. In addition, as the world is moving towards attaining the United Nations Sustainable Development (UNSD) goals by 2030, research aligned with any of the 17 UNSD goals is considered promising (Lafont-Torio et al., 2023). Therefore, in this article, an attempt was made to examine the renewable oils that are considered suitable replacements for fossil fuels.

Liquid fuels such as diesel, petrol, and kerosene, which are derived from fossil fuels, are often used in various power generation and heating systems, including internal combustion (IC) engines and boilers (Das et al., 2022). However, petro-diesel derived from fossil fuels increases environmental air pollution during combustion, which is considered a serious challenge to regulate (Kozina et al., 2020). To overcome the issue, edible and non-edible oils derived from different sources of plant seeds, fruits, nuts, leaves, wood, microbial and so forth are considered promising (Torres-García et al., 2020). Renewable oils are successfully utilised as second-grade fuel in different boilers, and first-grade bio-oils, also known as biofuel, are utilised in IC engines (Kolakoti et al., 2023a,b). In addition, bio-oils have applications in chemical synthesis, coke production, biochar production, chemical feedstocks, and so forth (Hu and Gholizadeh, 2020).

The suitability of renewable oils for different applications depends purely on their physical, chemical and thermal properties, and these characteristics play a predominant role in specific applications. For instance, the viscosity index and the direct usage of high-viscous oil are restricted in IC engines due to fuel atomization,

* Corresponding author.
E-mail addresses: aditya.kolakoti@gmail.com, akolakoti@imu.ac.in (A. Kolakoti).
Peer review under the responsibility of Editorial Office of Petroleum Research.

https://doi.org/10.1016/j.ptlrs.2025.02.009

fuel injection and clogging problems (Kolakoti and Appa Rao, 2020). In contrast, highly viscous oils are more likely to be used as bio-coolants in metal removal operations (Katam et al., 2023). Similarly, oils with poor heating value are less recommended for heating systems (Hu and Gholizadeh, 2020). This illustrates that the characterization of oils is crucial in determining their applications. However, the process of characterization involves the use of sophisticated instruments, and the experiments are time-consuming and costly. To address these challenges, artificial intelligence (AI) and machine learning (ML) techniques are gaining widespread attention, and these tools have exceptional problem-solving capabilities. They are able to solve non-linear problems with high accuracy, which helps in achieving sustainable development goals (Kolakoti et al., 2023a,b).

Zhang et al. (2022) reveal that it is crucial to predict the significant properties in bio-oils with high accuracy for their applications, and the biomass composition and pyrolysis conditions greatly influence the prediction outcomes. Therefore, the authors proposed a machine learning (ML) method to examine the different biomass compositions and pyrolysis conditions. Another investigation by Yang et al. (2022) estimated the bio-oil yield and oxygen content with the utilisation of the ML tool. The authors utilised a random forest algorithm for their predictions, and the pyrolysis conditions and biomass characteristics were used as input conditions to the algorithm. The authors conclude that two models (Ultimate-O and Proximate-Y) predict the oxygen content and bio-oil yield with high accuracy (0.895 R^2 and 0.925 R^2).

Chen et al. (2018) also proposed the artificial neural network (ANN) and support vector machine (SVM) algorithms to predict the bio-oil heating value and product distribution. For this purpose, nearly 200 correlated bio-mass samples were collected and trained to the algorithm to ensure the applicability of the proposed models. The modelling results reveal that ANN and SVM models perform reasonably well in estimating the pyrolytic bio-oil yield and heating value. The authors conclude that the proposed prediction models can be used as a good reference for studying biomass fast pyrolysis. In a related study, Leng et al. (2023) investigated the nitrogen heterocyclic compounds in bio-oil obtained from biomass pyrolysis using ML techniques of random forest and gradient boosting regression algorithms. The nitrogen heterocyclic compounds in bio-oils result in NOx formation when they are used as fuel in IC engines. In addition, prediction and control of these compounds are challenging in the pyrolysis process due to the complex chemical reactions. Therefore, the authors proposed ML techniques to emphasize the significant role of ML in providing valuable information for guiding experimental studies and addressing challenges associated with nitrogen heterocyclic compounds in bio-oils.

In recent times, image processing (IP) techniques have emerged as a popular tool for prediction and have been successfully used in various applications, including biomedical, fraud detection, flame characteristics and prediction properties (Jahanbakhshi et al., 2021; Kolakoti, 2024). By implementing the IP tools can reduce the overall experimental cost of examining properties. For instance, examination of significant fuel/oil properties for different combustion purposes. Caponi et al. (2023) applied the IP technique to predict the viscosity of hydrocarbon fuel (crude oil) present in heavy oil reservoirs in California, USA. Similarly, Kolakoti and Chandramouli (2023) investigated key fuel oil properties from different edible and non-edible oils using the IP tool. Therefore, with the utilisation of IP tools can regulate the time-intensive nature of expensive experimental procedures, which are considered as a limitation to the traditional methods. The IP works on signal processing that focuses on image manipulation, analysis, enhancement, or extracting significant information from the images. IP techniques are popularly used in different predictions, and a few research articles (Caponi et al., 2023; Kolakoti and Chandramouli, 2023) are available from the literature to examine the significant properties of renewable oils.

In this study, an attempt was made to examine the edible and non-edible oils significant properties based on the pixel information. The pixel information is extracted using two distant techniques of RGB (red, green, blue) and HSV (hue, saturation, value). These techniques offer valuable insights into the composition and characteristics of the biomass oils. The colour of biomass oil comes from different chemical compounds called chromophores, and these compounds affect the oil's properties. By analysing the composition of these chromophores, we can determine important characteristics of the oil. Our approach aims to contribute to the growing body of knowledge in this field, providing a unique perspective on renewable oil properties through image processing techniques.

2. Materials and methods

2.1. Selection of renewable oils

In this study, eleven renewable oils from both edible and non-edible categories were selected based on their availability and common use in various applications. The applications such as oils for biofuel production which are mostly obtained from non-edible categories like mahua, neem, castor, algae and waste cooking oil. Similarly, for bio-lubricant and bio-coolant for heat removal applications, different high and low viscous oils are used. The scientific/botanical names of these oils, along with their other common names, are presented in Table 1. Most of the non-edible oils were sourced from different locations, including Aaraku Valley in Visakhapatnam and the Thonam tribal market in Salur, while edible oils were obtained from nearby oil mills. The raw renewable oils underwent a water-washing and filtration process to remove suspended particles and any sediments using Whatman filter papers. After filtration, the oils were heated to the boiling point of water to eliminate water content and were then stored in airtight glass beakers. These oils were later subjected to testing for their significant physical, chemical, and thermal characteristics.

2.2. Experimental procedure

A set of renewable oil samples was sent for significant properties characterization, and the remaining oil was used for image processing (IP). To conduct this, we took a 250 ml clean and transparent glass beaker and poured 100 ml of clean and filtered oil into it. The glass beaker was positioned between two adjustable electrical fluorescent lamps (9-W capacity), as illustrated in Fig. 1. A digital camera (25 megapixels, Canon Make) was precisely placed above the glass beaker to capture surface images of the required

Table 1
Renewable oils with their scientific and common names.

S.No	Scientific Name	Common Name
1	Helianthus Annuus	Sunflower Oil
2	Chlorophyta	Green Algae Oil
3	Cocos Nucifera	Coconut Oil
4	Glycine Soja	Soybean Oil
5	Gossypium Herbaceum	Cotton Seed Oil
6	Azadirachta Indica Juss	Neem Oil
7	Madhuca Indica	Mahua Oil
8	Olea Europaea L	Olive Oil
9	Oryza Sativa L	Rice bran Oil
10	Ricinus Communis	Castor Oil
11	Waste-Cooking Oil (Elaeis Guineensis)	Waste-Cooking Oil (Palm Oil)

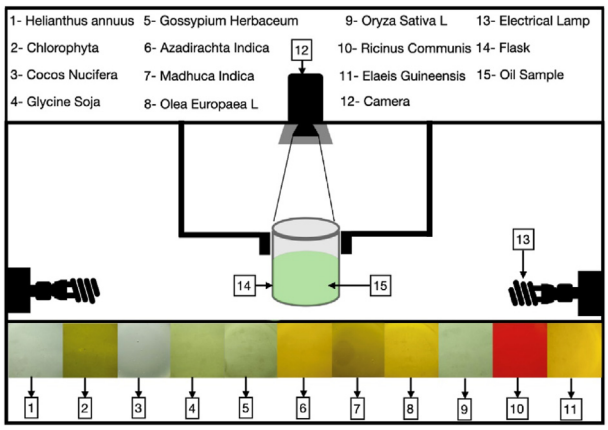


Fig. 1. Image Processing experimental setup.

renewable oil samples. Sensitivity to lighting conditions is a limitation of IP. For instance, if the camera is improperly positioned above the oil surface while capturing an image, it may capture shadows influenced by the light. These shadows could distort the pixels' data, which in turn could corrupt the data and ultimately affect the dataset, resulting in reduced accuracy. To avoid this, we place the camera parallel to the oil surface, where minimizing the effect of these shadows. In this way, we prevent light reflections from the lamps by adjusting their position and the camera's direction. To enhance accuracy, we captured around 150 surface images for each renewable oil sample under ambient conditions. These captured images are adequate to achieve reliable predictions for our specific research goal. All captured images were saved on a computer in the High Efficiency Image File (HEIF) format. HEIF can store the necessary data in a 10-bit format, as opposed to the 8-bit format utilised by the Joint Photographic Experts Group (JPEG). This means that HEIF files contain four times as many colours and tonal information as JPEG files (Yang et al., 2022). Due to this benefit, HEIF is chosen over other file formats. This procedure was followed for all oils to capture their surface images. Once the surface image capture was complete, we extracted significant information from the captured images using RGB and HSV techniques. This pixel metadata will be correlated to the results of the experimental oil properties.

2.3. Red Green Blue model

The RGB is an additive colour model that generates a diverse spectrum of colours by combining the light of varying intensities in Red, Green, and Blue. It finds widespread applications in various domains, especially in digital imaging, owing to its versatility in the representation and manipulation of colours (Zhang et al., 2022; Katam et al., 2024). RGB can be considered a fundamental model for encoding and displaying information about colours. Most electronic display instruments like laptops, computer monitors, television screens, smart mobile phones, digital display panels, and digital cameras make use of RGB to produce a wide range of colour spectrums accurately for better visibility to the human eye. In general, an 8-bit scale is used to express the RGB values, which have a 0 to 255 range, allowing 256 possible intensity levels for each colour channel (Su, 2020). For instance, pure red, pure green and pure blue intensities can be represented as (255, 0, 0), (0, 255, 0), and (0,0, 255), respectively. By combining these three RGB channels in varying proportions can develop a wide spectrum of colours (Equations 1–3). In our experimental study, the surface image of a given oil sample is captured and saved in a HEIF file format. The image is essentially a matrix of pixels, and each pixel is assigned to an RGB value. In this process, the relevant pixel data from the oil sample is associated with an RGB value, and this significant pixel information is extracted using Matlab software (R2023b).

$$Red = Red \,/\, (Red + Green + Blue) \tag{1}$$

$$Green = Green \,/\, (Red + Green + Blue) \tag{2}$$

$$Blue = Blue \,/\, (Red + Green + Blue) \tag{3}$$

2.4. Hue Saturation Value model

In the RGB method, the representation of colour is by mixing of red, green, and blue channels, and each pixel is a combination of these three colours. Similar to RGB, the HSV method in IP operates by differentiating the colours in the captured images based on their perceptual features. The perceptual features are hue (H), saturation (S), and value (V). The hue is a type of colour similar to red, green or blue, and it represents the dominant light wavelength in the image, which helps identify the basic colour in the object (Giuliani, 2022). The range of hue can be represented as 0–360° or 0 to 1 (normalized). The saturation (S) will represent the intensity or vividness of colour in the image, and high saturation colours are termed vibrant, pure, and low saturation colours are termed muted or pastel. This saturation will help classify how vibrant or pastel the colours are in the captured images (Giuliani, 2022). Finally, the value (V) represents the intensity or brightness of colours in the image, which helps identify how light or dark a colour is. Higher values represent the bright colours, whereas lower values represent the dark colours. The ranges of saturation and value are 0%–100% or 0 to 1 (normalization). The HSV techniques are popularly used in colour-based IP tasks like object recognition, image segmentation, image editing, and virtual and augmented reality (Flores-Vidal et al., 2022). The working mechanism of HSV can be illustrated in Figs. 2–4. The surface image of a given oil sample consists of a number of pixels, and each pixel is a combination of H, S, and V values. Fig. 2 represents the combination of H, S, and V values extracted from a given oil sample surface image. It is observed from Fig. 2 that H, S, and V represent their individual channels of hue, saturation and value.

In Fig. 3, the hue represents the colours present in the given oil sample, saturation indicates the intensity of colour, and the value indicates the brightness of the oil sample. These values are extracted from each pixel. Fig. 4 represents the distribution of total pixels with their corresponding HSV values for a given oil sample. In this way, the significant pixel information from the captured surface images of oils is utilised for their property analysis. Both techniques are unique in their specific applications as RGB excels at capturing subtle variations in the primary colour intensities, while HSV provides better insights into colour perception related to oil composition and viscosity.

2.5. Mathematical relations

The following mathematical relations are utilised for the analysis of pixel information from the surface images of captured oil samples. Equation (4) represents the mean (μ) values, denoting the sum of pixel values divided by the number of pixels (N). The standard deviation (SD) is presented in Equation (5), indicating the data points dispersion around the mean and the availability of variation from the mean pixel value. Equation (6) represents the

Fig. 2. Individual channels of a given surface image.

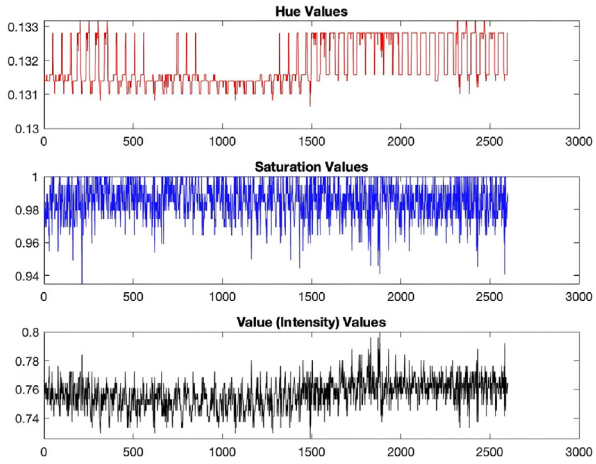


Fig. 3. Each channel intensity values.

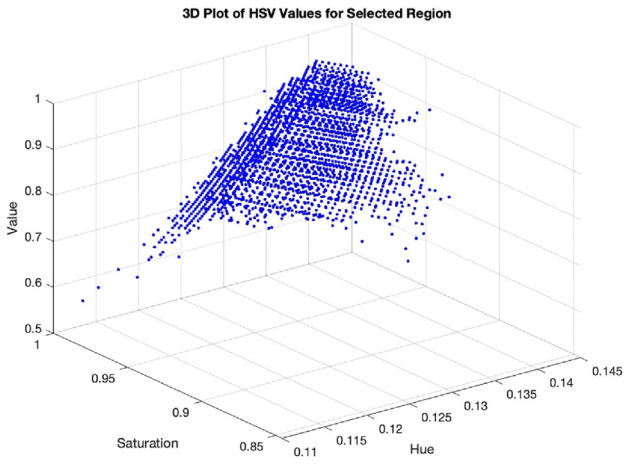


Fig. 4. Total pixels with their HSV representation.

Pearson correlation (r), which gives the relation between two similar colour pixels.

$$\mu = \frac{\sum x_i}{N} \tag{4}$$

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})}{n-1}} \tag{5}$$

$$r = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}} \tag{6}$$

3. Results & discussions

3.1. Significant oil properties characterization

The experimental characterization of significant oil properties, including physical, thermal, and chemical aspects, is considered premium, especially when the same oil sample is utilised for different property tests. Utilising the IP tools allows us to estimate various properties based on their pixel information (Caponi et al., 2023; Kolakoti and Chandramouli, 2023). For this purpose, different edible and non-edible oils, listed in Table 1, were selected for both experimental and IP analyses. Initially, all the raw oils are filtered, and captured their surface images for IP analysis. Simultaneously, the same raw oils are tested for their properties such as flash point (I), fire point (II), density (III), cloud point (IV), pour point (V) and viscosity (VI) experimentally as per the ASTM standards, and their results are presented in Table 2. The obtained experimental results of eleven oil sample are mapped to the RGB and HSV metadata. From the oil property analysis, the highest and lowest kinematic viscosities were observed for the Ricinus communis, Madhuca indica, which are measured with a redwood viscometer by following ASTM-D445 standards. Viscosity can be considered one of the significant properties of oil that helps estimate its flow behaviour. For instance, low viscous oils can be directly used as fuel in CI engines due to easier combustion atomization (Kolakoti and Appa Rao, 2020; Kolakoti et al., 2023a,b). Conversely, high viscosity may lead to poor atomization and fuel injector clogging problems (Kolakoti and Appa Rao, 2020) and is more suitable for bio-coolant applications (Katam et al., 2024). Compared to petrochemical oils, renewable oils are reactive to atmospheric conditions when they are exposed. This is due to changes in the composition of fatty acids, which results in the alteration of properties (Atabani et al., 2013; Kolakoti and Appa Rao, 2019). Therefore, the cold flow properties are examined for oils as per ASTM-D97 standards, and the results of cloud and pour points reveal that madhuca indica is recorded as the highest. Finally, the density and ease of storing and transportation properties of flash and fire points are examined as per the ASTM-D4052 and D92 standards, and it is observed that these oils have higher values than the petrochemical oils, which can be considered a good indication for easy storing and transportation. Finally, the measured significant physical properties of each oil sample are mapped with the digital pixel metadata. This metadata is obtained from the surface image by extracting pixel information using two distinct colour models (RGB and HSV).

Table 2
Significant oil properties of edible and non-edible oils.

S.No	List of oils	I	II	III	IV	V	VI
1	Helianthus Annuus	329	340	925.88	−6.7	−8.3	53.41
2	Chlorophyta	124	135	930.12	−7	−9	29.66
3	Cocos Nucifera	213	229	926.24	−8	7	31.32
4	Glycine Soja	332	364	919.31	0.7	−10	37.42
5	Gossypium Herbaceum	247	253	918.23	2	−6	32.53
6	Azadirachta Indica Juss	309	321	923.76	12	7	212.66
7	Madhuca Indica	236	249	924.89	10	7	25.81
8	Olea Europaea L	329	341	920	7	−1	42.31
9	Oryza Sativa L	333	375	923.46	9	−3	49.87
10	Ricinus Communis	331	350	962.87	6	4	261.21
11	Waste-cooking oils (Elaeis Guineensis)	335	348	916.73	6	1	41.26

I = Flash Point (^{0}C); II = Fire Point (^{0}C); III = Density (kg/m^3); IV = Cloud Point (^{0}C); V = Pour Point (^{0}C); VI = Viscosity (cST@40^{0}C).

3.2. RGB additive colour model analysis

The default mode of storing a colour space in digital images is RGB, which is widely used in digital image analysis applications. The RGB colour model is also known as the additive colour model, and due to its compatibility, any other colour can be produced by mixing three base colours, and each specific colour is categorized as red, green and blue, as shown in Equations (1–3). In this study, the surface images of oil samples are stored in the HEIF file format and analysed for extraction of significant RGB information using Matlab software (R2023b). The software extracts the RGB information of all the oil samples (11 oils × 150 surface images = 1650 surface images) in terms of mean, mode, maximum, minimum, and standard deviation (SD) as shown in Fig. 5, and the average results are presented in Table 3. Additionally, the RGB intensity of all the oil samples in-terms of 3D surface plots are presented in Fig. 5a–k. The axis X and Y defines the coordinates of each pixel in an oil sample image, and the Z direction represents the pixel intensities of the oil (Kolakoti and Chandramouli, 2023). From the plots, it is evident that there is a similarity in red, green and blue channel surface plots. However, the intensity values are varied. Fig. 5c, d, h, i, and j indicates that the density and viscosity layers of the oil samples are thin, resulting in better visibility of the background white surface. On the other hand, oils belong to higher density and viscosity; the visibility of the background white surface area is limited, resulting in variation of colour channel values as shown in Fig. 5a, b, e, f, g, and k.

3.3. HSV perceptual colour model analysis

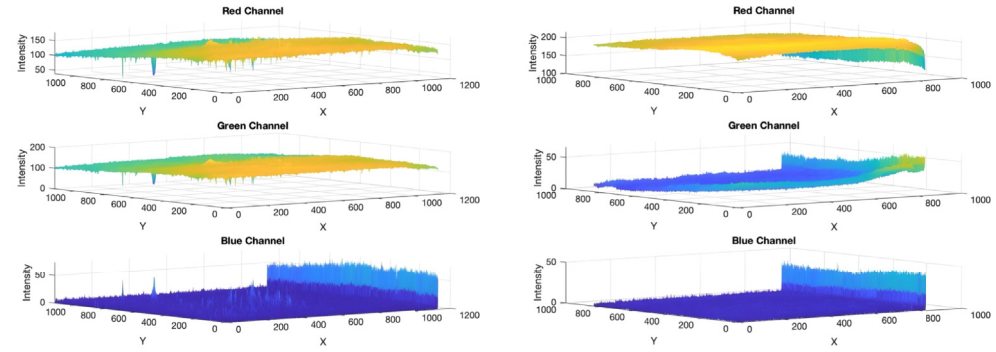
Another widely used method for extracting pixel information from captured images is the HSV technique. In this approach, the 1650 surface images of all the oil samples saved in HEIF file format are analysed to extract the oil sample pixel metadata. The extracted pixel information includes mean, maximum, minimum, mode and standard deviation (SD), as shown in Fig. 6. The average results of 1650 surface images are presented in Table 4. The 3D surface channels of HSV for all the tested oil samples are shown in Fig. 6a–k. In each image, the X and Y axis represent the coordinates of the oil sample in terms of pixel count, while the Z axis denotes the intensities of hue, saturation and value channels, respectively (Kolakoti and Chandramouli, 2023). The hue contains different shades of colour, starting from red, yellow, green, cyan, blue, and black, followed by red. In general, 0 to 1 is the intensity values ranges for HSV and the maximum intensity of saturation is witnessed for Chlorophyta, Ricinus Communis, Madhuca Indica, Azadirachta Indica Juss, Olea Europaea L and Elaeis Guineensis as shown in Fig. 6a, b, e, f, g, and k. This may be due to the presence of high pigments, which absorb a certain wavelength of light and reflect others, thereby giving them their colour (Giuliani, 2022). On the other hand, the lowest saturation values are observed for Cocos Nucifera, Gossypium Herbaceum, Oryza Sativa L, Glycine Soja, and Helianthus Annuus, as shown in Fig. 6c, d, h, i, and j. Finally, the value channel indicates the brightness of the colour pixel and the maximum value is recorded for Cocos Nucifera and Madhuca Indica as 1, and the minimum value is recorded for Chlorophyta as 0.145, as shown in Table 4.

3.4. Assignment of RGB and HSV with Fiji

The accuracy of the developed RGB and HSV models pixel intensity in terms of min, max, mode, mean and standard deviation (SD) are validated with Fiji open-source software (Dumont et al., 2021). Fiji is popularly used for analysing complex biological images (Cayuela López et al., 2022; Deshpande et al., 2021) and can perform with better accuracy in analysing microscopy images. Fiji is based on a popular IP program called ImageJ and has a wide range of image analysis tools, which include image filtration, visualization, segmentation and quantification, which enables the user to perform complex IP tasks more easily (Ahammer et al., 2023). Moreover, Fiji is made available freely as an open source. Therefore, significant pixel information obtained from RGB and HSV are analysed in Fiji. Tables 5 and 6 represent the experimental average values of mean, mode and SD of RGB and HSV compared to the Fiji results. This shows that the experimental and Fiji results are in good agreement with high accuracy. However, a minor error was observed in the standard deviation parameter due to the handling of outliers in the pixel information and algorithms. The Fiji obtained RGB and HSV mean, mode, and SD results for waste cooking oil are shown in Figs. 7 and 8, which can be compared with Figs. 5 k and Fig. 6 k, respectively.

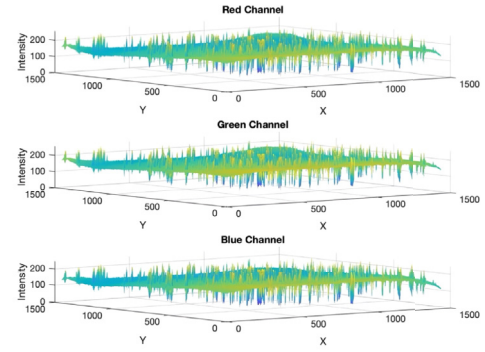
3.5. Validation

The accuracy of the developed dataset (RGB data values 34,650 and HSV data values 34,650) from 1650 surface images is further validated with the supervised machine learning algorithms of Neural Network classification and TreeBagger. The tree-bagger algorithm consists of many decision trees, and each tree is utilised during the training process (Yagmur et al., 2023). These decision trees help make more accurate decisions by combining their individual opinions from many different decision trees (Yagmur et al., 2023). High prediction accuracy can be achieved with this tree-bagger algorithm as it works similarly to a group of experts (multiple decision trees) analysing a particular problem to make a better decision together. In this study, the developed datasets of RGB and HSV are trained in the treebagger algorithm, and the obtained results are shown in Figs. 9 and 10. Each plot represents the Receiver

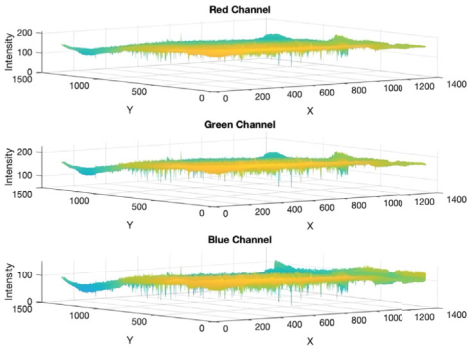


(a) Chlorophyta

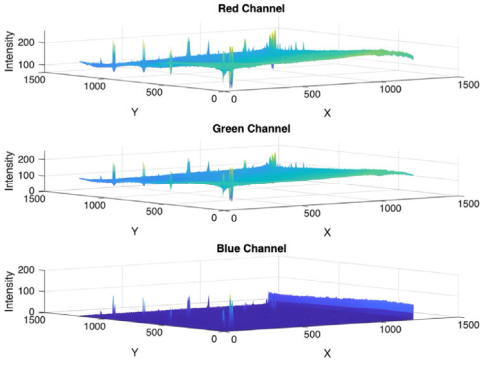
(b) Ricinus Communis



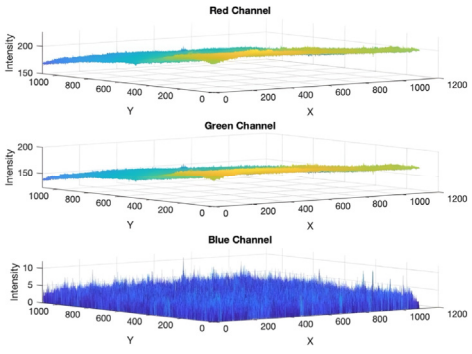
(c) Cocos Nucifera



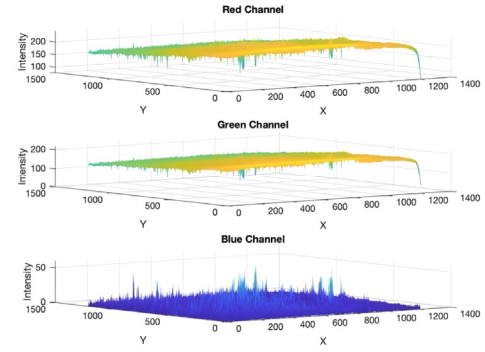
(d) Gossypium Herbaceum



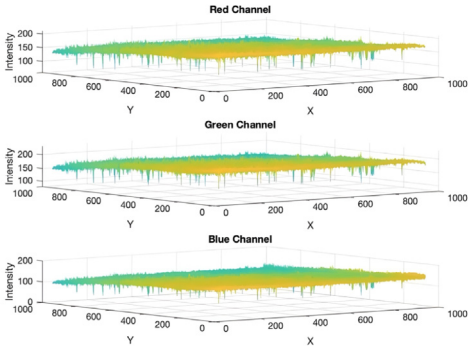
(e) Madhuca Indica



(f) Azadirachta Indica Juss

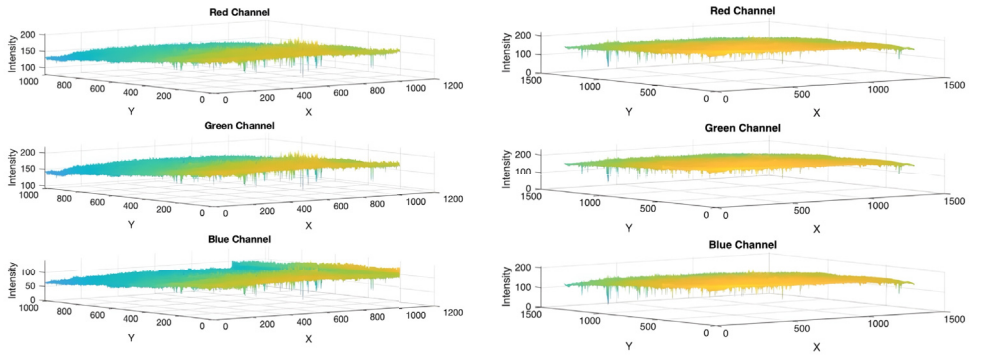


(g) Olea Europaea L

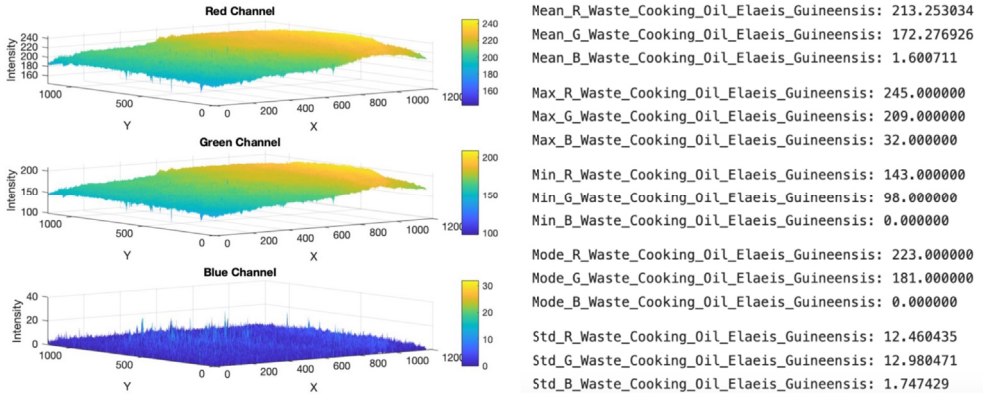


(h) Oryza Sativa L

Fig. 5. 3D Surface plots of RGB intensity in oil samples.



(i) Glycine Soja (j) Helianthus Annuus



(k) Waste cooking oil

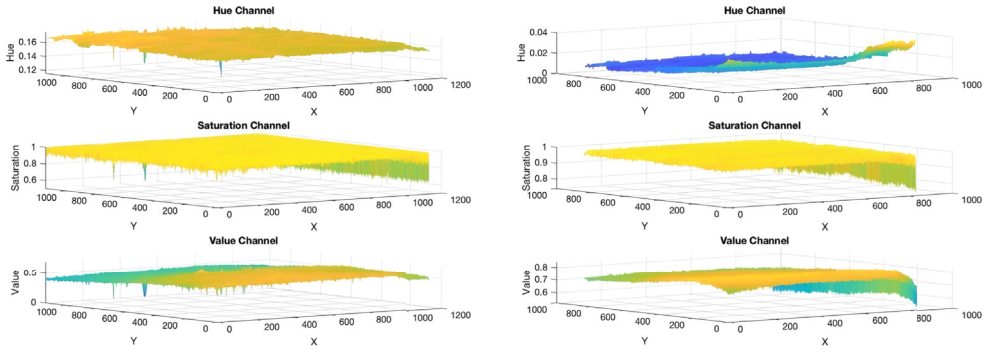
Fig. 5. (*continued*).

Table 3
The average results of RGB intensity.

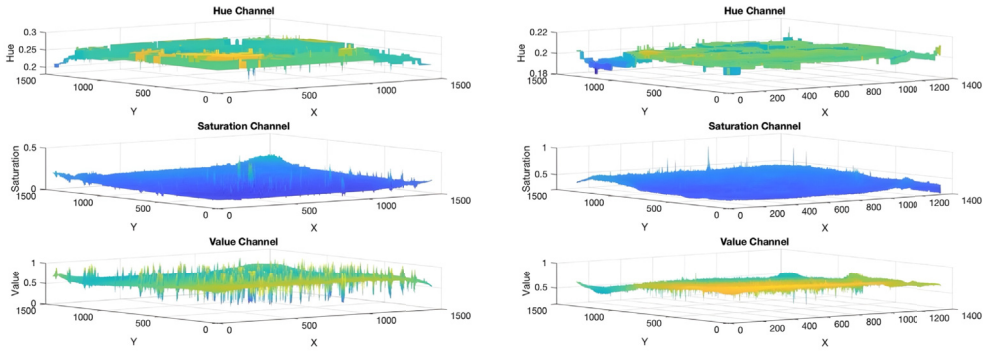
S.No			Helianthus annuus	Chlorophyta	Cocos Nucifera	Glycine Soja	Gossypium Herbaceum	Azadirachta indica	Madhuca Indica	Olea Europaea L	Oryza Sativa L	Ricinus Communis	Elaeis Guineensis
1	Mean	R	172	134	159	157	152	191	151	198	160.88	199	213.25
2		G	185	130	169	171	166	159	128	170	179.33	17.75	172.27
3		B	160	1.40	147	93	98	0.97	1.08	1.94	126.80	1.56	1.60
4	Max	R	220	177	254	204	210	226	255	243	210	215	245
5		G	232	171	255	218	226	193	255	218	230	66	209
6		B	208	75	245	146	155	12	169	60	177	47	32
7	Min	R	36	37	35	80	34	155	67	78	59	117	143
8		G	45	30	41	94	48	121	35	39	80	2	98
9		B	16	0	24	17	0	0	0	0	24	0	0
10	Mode	R	180	148	167	146	156	190	160	203	171	205	223
11		G	193	144	179	162	171	157	136	173	192	12	181
12		B	166	0	158	80	104	0	0	0	138	1	0
13	SD	R	14	14.57	18	12	19	11.89	19.64	14.92	11.74	6.63	12.46
14		G	14	13.37	17	13	20	12.26	19.83	16.06	12.90	7.91	12.98
15		B	15	2.65	20	14	18	1.16	2.88	2.57	12.60	1.88	1.74

Operating Characteristics (ROC) curve between the true and false positive rates on the axis. The curves near to 1 at the upper left corner of the true positive rate indicate that the trained dataset is highly accurate (91.9% for RGB and 95.3% for HSV) and possesses maximum sensitivity and specificity.

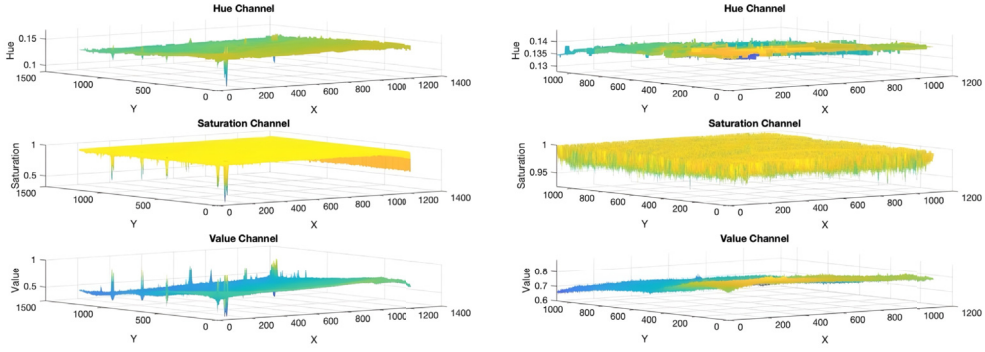
Another machine learning algorithm used in this study is Neural Network (NN) classification. The NN classification algorithm works similarly to the human brain, is powerful in handling complex patterns in data, and is widely used in image recognition (Kolakoti et al., 2023a,b). It consists of three layers named as input, output and hidden, and an activation function is utilised to learn the complex relationships in the input and output data (Ashtiani et al., 2022; Kolakoti et al., 2023a,b). The results obtained from the NN classification algorithm are shown in parallel coordinates in Figs. 11 and 12. For ease of understanding, the parallel coordinate plot of one oil sample (Chlorophyta) is considered and the X axis represents the digital and experimental oil properties, and the Y axis represents their ranges. Each line represents the data point that connects to its values on the axis, and the connected line segments to these data points are referred to as polylines. The axis can be scaled differently according to the range of each variable. Therefore, on the X-axis, oil properties and statistical parameters of mean,



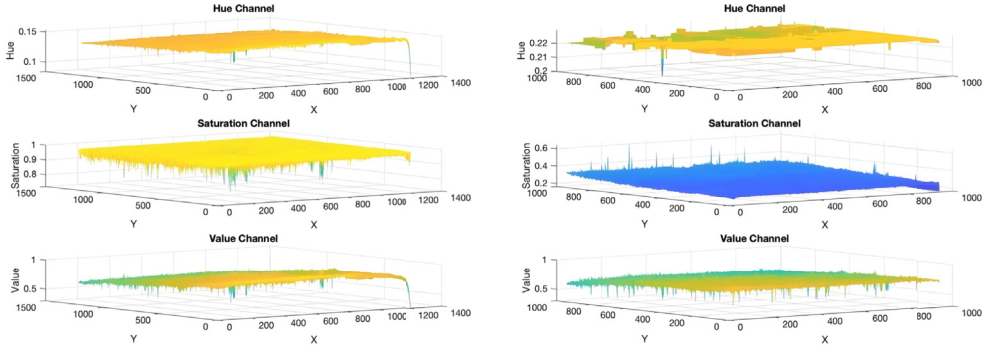
(a) Chlorophyta (b) Ricinus Communis



(c) Cocos Nucifera (d) Gossypium Herbaceum

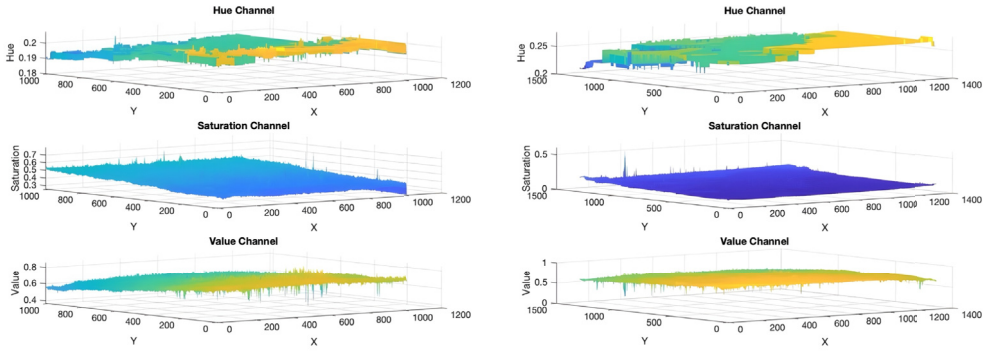


(e) Madhuca Indica (f) Azadirachta Indica Juss



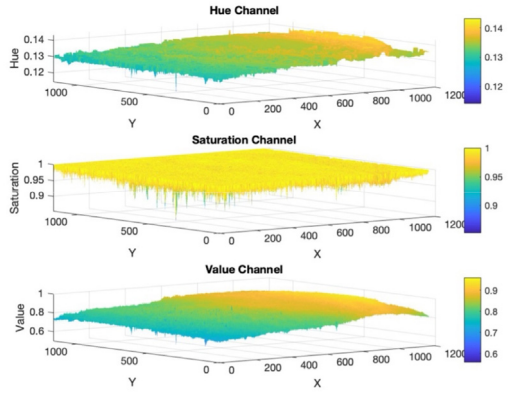
(g) Olea Europaea L (h) Oryza Sativa L

Fig. 6. 3D Surface plots of HSV channels in oil samples.



(i) Glycine Soja

(j) Helianthus Annuus




```
Mean_H_Waste_Cooking_Oil_Elaeis_Guineensis: 0.134278
Mean_S_Waste_Cooking_Oil_Elaeis_Guineensis: 0.992603
Mean_V_Waste_Cooking_Oil_Elaeis_Guineensis: 0.836286

Max_H_Waste_Cooking_Oil_Elaeis_Guineensis: 0.143275
Max_S_Waste_Cooking_Oil_Elaeis_Guineensis: 1.000000
Max_V_Waste_Cooking_Oil_Elaeis_Guineensis: 0.960784

Min_H_Waste_Cooking_Oil_Elaeis_Guineensis: 0.114219
Min_S_Waste_Cooking_Oil_Elaeis_Guineensis: 0.853211
Min_V_Waste_Cooking_Oil_Elaeis_Guineensis: 0.560784

Mode_H_Waste_Cooking_Oil_Elaeis_Guineensis: 0.135135
Mode_S_Waste_Cooking_Oil_Elaeis_Guineensis: 1.000000
Mode_V_Waste_Cooking_Oil_Elaeis_Guineensis: 0.874510

Std_H_Waste_Cooking_Oil_Elaeis_Guineensis: 0.002568
Std_S_Waste_Cooking_Oil_Elaeis_Guineensis: 0.008027
Std_V_Waste_Cooking_Oil_Elaeis_Guineensis: 0.048864
```

(k) Waste cooking oil

Fig. 6. (*continued*).

Table 4
The average results of HSV intensity.

S.No			Helianthus annuus	Chlorophyta	Cocos Nucifera	Glycine Soja	Gossypium Herbaceum	Azadirachta indica	Madhuca Indica	Olea Europaea L	Oryza Sativa L	Ricinus Communis	Elaeis Guineensis
1	Mean	H	0.2531	0.1621	0.2442	0.1973	0.2017	0.1383	0.1405	0.1427	0.2251	0.0135	0.1342
2		S	0.1344	0.9896	0.1301	0.4553	0.4109	0.9949	0.9928	0.9904	0.2942	0.992	0.9926
3		V	0.7278	0.5273	0.6638	0.6724	0.6533	0.7504	0.5938	0.7793	0.7032	0.7804	0.8362
4	Max	H	0.2754	0.1733	0.287	0.2133	0.2152	0.1449	0.1666	0.1522	0.2403	0.0429	0.1423
5		S	0.6444	1	0.5294	0.8191	1	1	1	1	0.7	1	1
6		V	0.9098	0.6941	1	0.8549	0.8862	0.8862	1	0.9529	0.9019	0.8431	0.9607
7	Min	H	0.2028	0.1129	0.1785	0.1812	0.1833	0.1272	0.0869	0.0833	0.1928	0.0018	0.1142
8		S	0.0982	0.4573	0.0392	0.1965	0.1807	0.9408	0.3372	0.7222	0.1263	0.6466	0.8532
9		V	0.1764	0.145	0.1607	0.3686	0.1882	0.6078	0.2627	0.3058	0.3137	0.4588	0.5607
10	Mode	H	0.25	0.161	0.2456	0.1962	0.2039	0.1368	0.1382	0.1419	0.2295	0.0147	0.1351
11		S	0.125	1	0.1073	0.5	0.4294	1	1	1	0.276	1	1
12		V	0.7568	0.5603	0.7019	0.6352	0.6705	0.745	0.6274	0.796	0.752	0.8039	0.8745
13	SD	H	0.0117	0.0025	0.0111	0.0032	0.004	0.0023	0.0039	0.0031	0.0046	0.0062	0.0025
14		S	0.0198	0.0205	0.0316	0.0423	0.0449	0.006	0.0177	0.0123	0.0223	0.012	0.008
15		V	0.0554	0.057	0.0702	0.0513	0.0798	0.0466	0.077	0.0585	0.0505	0.026	0.0488

Table 5
Comparison of RGB and Fiji results.

S.No		Experimental RGB results			Fiji results		
		R	G	B	R	G	B
1	Mean	213.25	172.27	1.600	213.25	172.28	1.60
2	Mode	223	181	0	223	181	0
3	SD	12.46	12.98	1.747	NaN	12.98	1.75

Table 6
Comparison of HSV and Fiji results.

S.No		Experimental HSV results			Fiji results		
		H	S	V	H	S	V
1	Mean	0.1342	0.996	0.836	0.135	0.987	0.834
2	Mode	0.135	1.0	0.874	0.136	1.0	0.871
3	SD	0.00256	0.00802	0.0488	0.00243	0.00968	0.0484

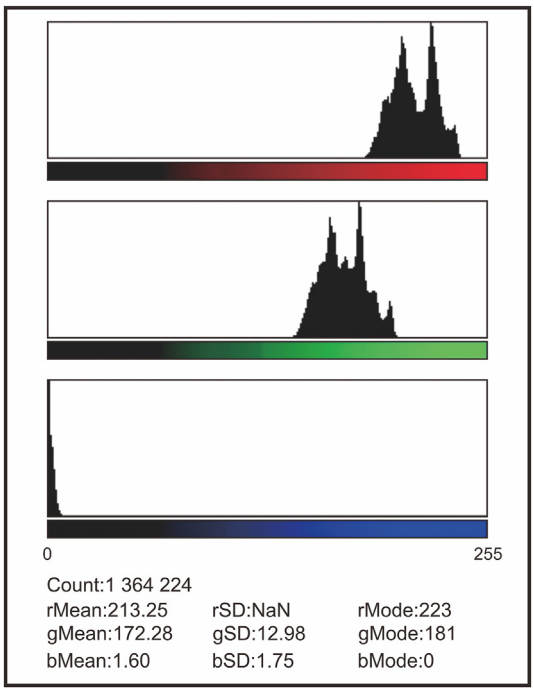


Fig. 7. Fiji results of RGB for waste cooking oil.

mode, max, min and SD are presented. The figures show that the variations in mean and standard deviations from 150 images of the oil sample are mapped with other parameters due to the changes in the range. However, no significant change in the ranges is observed for maximum, minimum, mode, and oil properties. The accuracy of the NN classification model for RGB and HSV is computed as 95.6% and 97.3%, respectively.

Finally, two different oils not used in this study are chosen to test the prediction accuracy of the developed dataset. For this purpose, Arachis Hypogaea (Groundnut Oil) and Jatropha Curcas (Jatropha Oil) are considered, and their surface images are captured as per the procedure illustrated in section 2.2. The significant pixel information regarding RGB and HSV are extracted, and the average statistical parameters of mean, max, min, mode and SD results of Arachis Hypogaea and Jatropha Curcas are presented in Table 7. The obtained pixel intensities of the two oil samples are correlated with the 11 oils dataset, and Table 8 illustrates the experimental oil properties versus RGB and HSV predicted oil properties. It is observed that RGB achieves 84.68% and 85.61% accuracy in the prediction of significant oil properties of Arachis Hypogaea and Jatropha Curcas. In contrast, HSV achieves 88.53% and 88.40% accuracy in predicting the same oil properties. The lower accuracy of the RGB model may be attributed to the oil's high saturation levels, which are better captured by the HSV model. HSV has the advantage of capturing multiple shades of colour, which is a limitation of RGB.

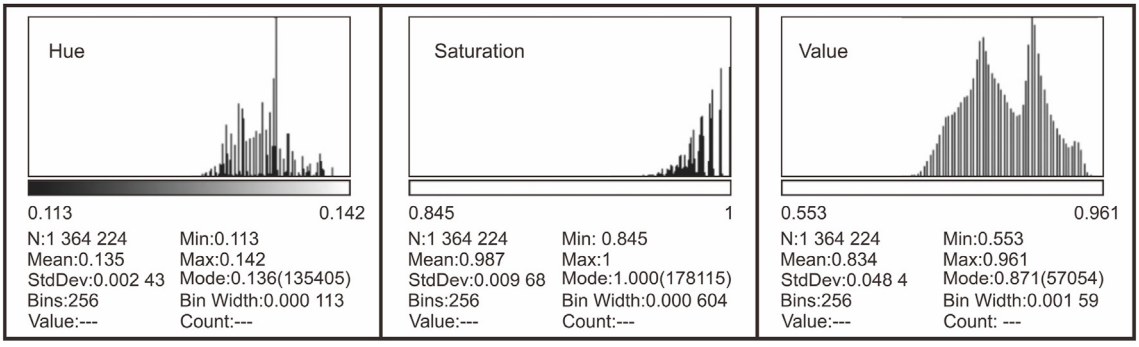


Fig. 8. Fiji results of HSV for waste cooking oil.

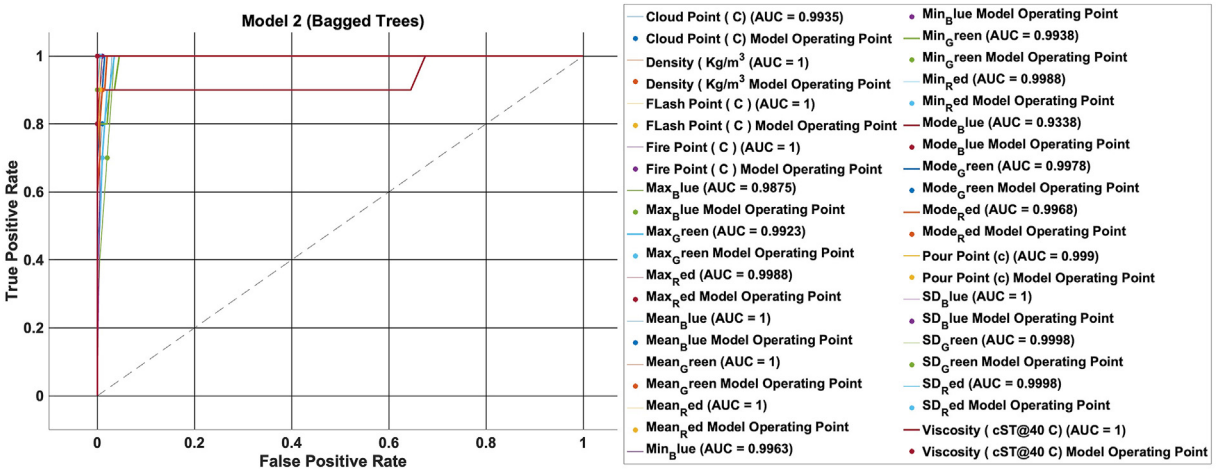


Fig. 9. Receiver Operating Characteristics curve for RGB dataset.

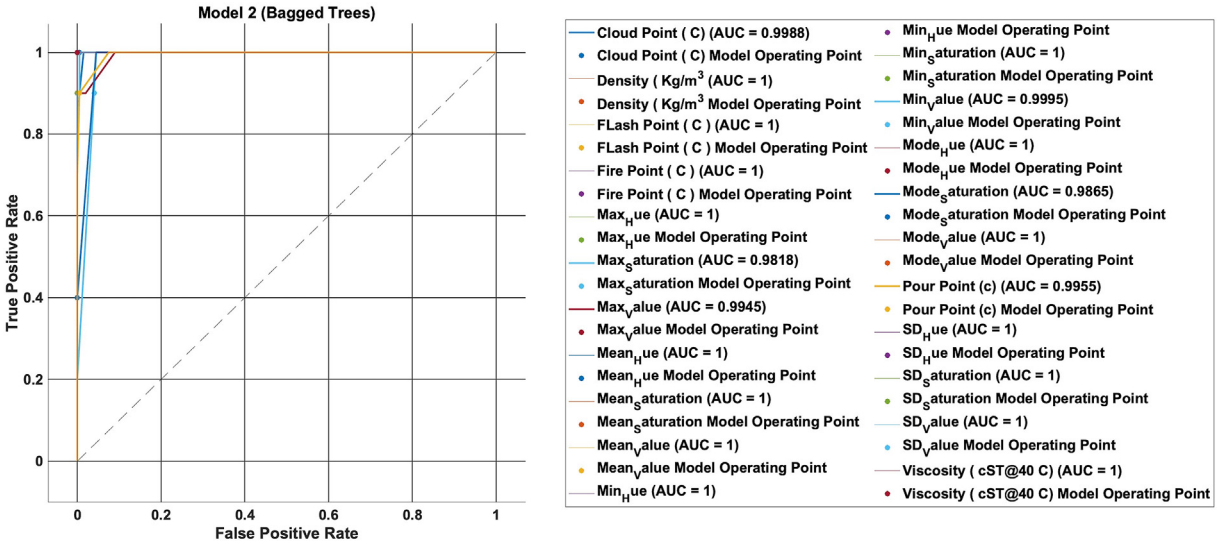


Fig. 10. Receiver Operating Characteristics Curve for HSV dataset.

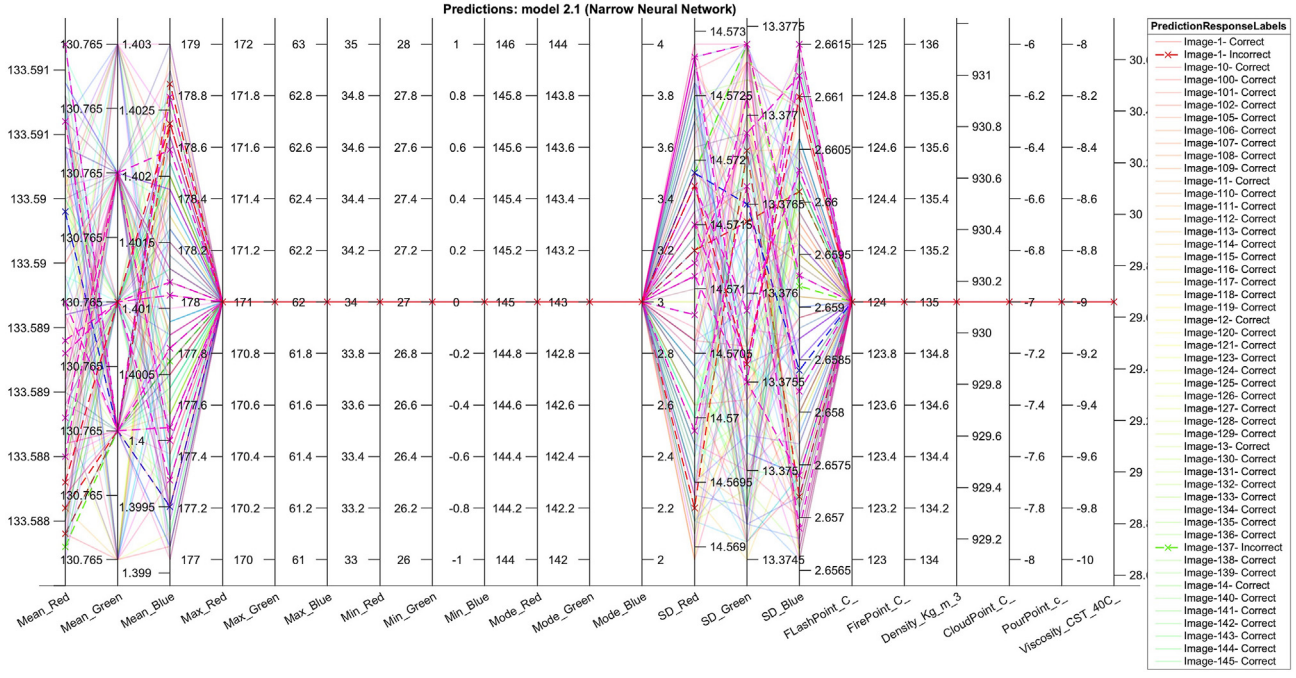


Fig. 11. Neural Network classification results for RGB dataset.

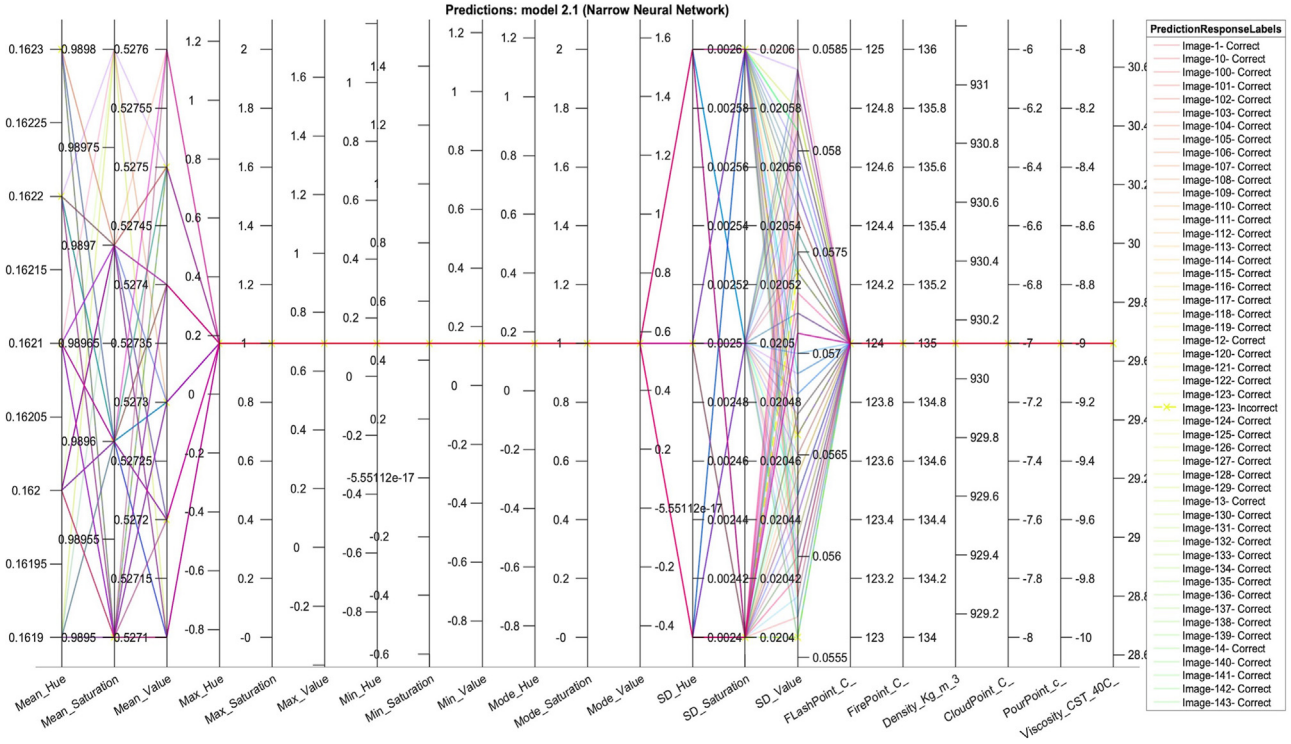


Fig. 12. Neural Network classification results for HSV dataset.

Table 7
RGB and HSV statistical parameters analysis.

S.No	Statistical Parameters	Colour Channels	Arachis Hypogaea	Jatropha Curcas
1	Mean	Hue	0.1552	0.0738
		Saturation	0.9935	0.9916
		Value	0.6994	0.8241
		Red	178.3484	210.1603
		Green	166.2341	94.1729
		Blue	1.1753	1.7361
2	Max	Hue	0.1679	0.0824
		Saturation	1.0000	1
		Value	0.8824	1
		Red	225.0000	255
		Green	215.0000	149
		Blue	103.0000	69
3	Min	Hue	0.1303	0.0460
		Saturation	0.4831	0.5886
		Value	0.4314	0.5725
		Red	110.0000	146
		Green	86.0000	44
		Blue	0.0000	0
4	Mode	Hue	0.1553	0.0714
		Saturation	1.0000	1
		Value	0.6941	0.8039
		Red	177.0000	205
		Green	164.0000	88
		Blue	0.0000	0
5	SD	Hue	0.0024	0.0023
		Saturation	0.0166	0.0184
		Value	0.0603	0.0284
		Red	15.3814	7.2647
		Green	15.2458	5.9185
		Blue	3.0080	3.1780

Table 8
Oil properties prediction.

S.No	Properties	Arachis Hypogaea			Jatropha Curcas		
		Exp	RGB	HSV	Exp	RGB	HSV
1	Flash Point (^{0}C)	331	327	330	305	307	307
2	Fire Point (^{0}C)	343	340	345	320	319	321
3	Density (kg/m^3)	924.47	920.0	921	943.41	942.0	941.0
4	Cloud Point (^{0}C)	3	1	5	2	3	2
5	Pour Point (^{0}C)	−5	−6	−5	6	4	2
6	Viscosity (cST@40^{0}C)	55.48	54.0	55	83.42	85.0	82.0
7	Accuracy		**84.68%**	**88.53%**		**85.61%**	**88.40%**

Finally, our experimental results are compared with the existing literature of Caponi et al. (2023) at Texas A&M University, which showed that a similar working principle was implemented, and they achieved high accuracy. T. Zhang et al. (2022) also achieved less than 20% relative error from their findings.

4. Conclusions

The renewable oils used in this study belong to edible and non-edible categories, and their essential oil properties like viscosity, density, flash point, fire point, cloud, and pour points are measured experimentally by following recommended ASTM standards. Around 150 surface images are captured for each edible and non-edible oil, and the significant pixel data information from the captured images is extracted using RGB and HSV methods. The RGB and HSV extract the pixel intensity from the digital images (1650) of oil samples and present the average results in terms of mean, mode, maximum, minimum, and standard deviation. The extracted pixel information from RGB and HSV methods are verified with Fiji open image processing software, and the results obtained from Fiji (mean, mode, and standard deviation) match the RGB and HSV results.

Two supervised machine learning algorithms, Neural Network Classification and TreeBagger, are utilised to validate the developed RGB and HSV datasets further. The neural network classification algorithm reveals that RGB and HSV datasets are 95.6% and 97.3% accuracy, and the treebagger shows that receiver operating characteristics curves for RGB and HSV are very close to one, indicating that the true positive rate is accurate and achieves 91.9% and 95.3% accuracy for RGB and HSV dataset. The developed digital pixel metadata from the edible and non-edible oils are correlated with the experimental oil properties to predict any renewable oil properties. Arachis Hypogaea and Jatropha Curcas oils, which are not in the experiment, are chosen to predict their oil properties using RGB and HSV datasets. Interestingly, the RGB method accurately predicts the Arachis Hypogaea and Jatropha Curcas oil properties with 84.68% and 85.61%, and HSV predicts 88.53% and 88.40% accurately. Therefore, these findings suggest that RGB and HSV image processing techniques offer a cost-effective and reliable means of predicting significant oil properties.

These techniques may be extended to the production process of biofuels, which are considered to be one of the promising fields in promoting sustainable development goals. However, sensitivity to light conditions and angle of view while capturing the surface images are the limitations that can be minimised with proper care and controlled atmosphere, and also expanding the dataset in future research could further enhance the model's predictive power and generalization across a wider range of oils.

CRediT authorship contribution statement

Aditya Kolakoti: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Investigation. **Ruthvik Chandramouli:** Software, Methodology, Formal analysis, Data curation.

Data availability

The datasets developed during the current study are available from the corresponding author upon reasonable request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ahammer, H., Reiss, M.A., Hackhofer, M., Andronache, I., Radulovic, M., Labra-Spröhnle, F., Jelinek, H.F., Comsystan, J., 2023. A collection of Fiji/ImageJ2 plugins for nonlinear and complexity analysis in 1D, 2D and 3D. PLoS One 18 (10), e0292217. https://doi.org/10.1371/journal.pone.0292217.

Ashtiani, F., Geers, A.J., Aflatouni, F., 2022. An on-chip photonic deep neural network for image classification. Nature 606, 501–506. https://doi.org/10.1038/s41586-022-04714-0.

Atabani, A., Silitonga, A., Ong, H., Mahlia, T., Masjuki, H., Badruddin, I.A., Fayaz, H., 2013. Non-edible vegetable oils: a critical evaluation of oil extraction, fatty acid compositions, biodiesel production, characteristics, engine performance and emissions production. Renew. Sustain. Energy Rev. 18, 211–245. https://doi.org/10.1016/j.rser.2012.10.013.

Caponi, M., Cox, A., Misra, S., 2023. Viscosity prediction using image processing and supervised learning. Fuel 339, 127320. https://doi.org/10.1016/j.fuel.2022.127320.

Cayuela López, A., Gómez-Pedrero, J.A., Blanco, A.M.O., Sorzano, C.O.S., 2022. CellTypeAnalyzer: a flexible Fiji/ImageJ plugin to classify cells according to user-defined criteria. Biological Imaging 2, e5. https://doi.org/10.1017/S2633903X22000058.

Chen, X., Zhang, H., Song, Y., Xiao, R., 2018. Prediction of product distribution and bio-oil heating value of biomass fast pyrolysis. Chemical Engineering and Processing - Process Intensification 130, 36–42. https://doi.org/10.1016/j.cep.2018.05.018.

Das, A.K., Sahu, S.K., Panda, A.K., 2022. Current status and prospects of alternate liquid transportation fuels in compression ignition engines: a critical review. Renew. Sustain. Energy Rev. 161, 112358. https://doi.org/10.1016/j.rser.2022.112358.

Deshpande, P., Gogia, N., Chimata, A.V., Singh, A., 2021. Unbiased automated quantitation of ROS signals in live retinal neurons of Drosophila using Fiji/ImageJ. Biotechniques 71 (2), 416–424. https://doi.org/10.2144/btn-2021-0006.

Dumont, L., Levacher, N., Schapman, D., Rives-Feraille, A., Moutard, L., Delessard, M., Saulnier, J., Rondanino, C., Rives, N., 2021. IHC_Tool: an open-source Fiji procedure for quantitative evaluation of cross sections of testicular explants. Reprod. Biol. 21 (2), 100507. https://doi.org/10.1016/j.repbio.2021.100507.

Flores-Vidal, P., Gómez, D., Castro, J., Montero, J., 2022. New aggregation approaches with HSV to color edge detection. Int. J. Comput. Intell. Syst. 15, 78. https://doi.org/10.1007/s44196-022-00137-x.

Giuliani, D., 2022. Metaheuristic algorithms applied to color image segmentation on HSV space. J. Imaging 8 (1), 6. https://doi.org/10.3390/jimaging8010006.

Hu, X., Gholizadeh, M., 2020. Progress of the applications of bio-oil. Renew. Sustain. Energy Rev. 134, 110124. https://doi.org/10.1016/j.rser.2020.110124.

Jahanbakhshi, A., Abbaspour-Gilandeh, Y., Heidarbeigi, K., Momeny, M., 2021. Detection of fraud in ginger powder using an automatic sorting system based on image processing technique and deep learning. Comput. Biol. Med. 136, 104764. https://doi.org/10.1016/j.compbiomed.2021.104764.

Katam, A.K., Mohanty, R.C., Kolakoti, A., 2023. The role of bio-based cutting fluids for sustainable manufacturing and machining processes: a holistic review. Mechanical Engineers for Society and Industry 3 (3), 166–180. https://doi.org/10.31603/mesi.10680.

Katam, A.K., Mohanty, R.C., Kolakoti, A., 2024. Experimental and image processing-based characterization of sustainable bio-coolant for metal removal operations. J. Braz. Soc. Mech. Sci. Eng. 46, 216. https://doi.org/10.1007/s40430-024-04770-9.
Kolakoti, A., 2024. Influence of magnetic field on renewable oil flame characteristics-an experimental and image processing analysis with bibliometric study. e-Prime - advances in Electrical Engineering. Electronics and Energy 10, 100776. https://doi.org/10.1016/j.prime.2024.100776.
Kolakoti, A., Appa Rao, B.V., 2019. Effect of fatty acid composition on the performance and emission characteristics of an IDI supercharged engine using neat palm biodiesel and coconut biodiesel as an additive. Biofuels 10 (5), 591–605. https://doi.org/10.1080/17597269.2017.1332293.
Kolakoti, A., Appa Rao, B.V., 2020. Relative testing of neat Jatropha methyl ester by preheating to viscosity saturation in IDI engine - an optimisation approach. Int. J. Automot. Mech. Eng. 17 (2), 8052–8066. https://doi.org/10.15282/ijame.17.2.2020.23.0604.
Kolakoti, A., Bobbili, P., Katakam, S., Geeri, S., Soliman, W.G., 2023a. Applications of artificial intelligence in sustainable energy development and utilization. In: Malik, S.C., Sinwar, D., Kumar, A., Gadde, S.R., Chatterjee, P., Bui, H. (Eds.), Computational Intelligence in Sustainable Reliability Engineering. Scrivener Publishing Wiley. https://doi.org/10.1002/9781119865421.ch6.
Kolakoti, A., Chandramouli, R., 2023. Prediction of significant oil properties using image processing based on RGB pixel intensity. Fuel 349, 128618. https://doi.org/10.1016/j.fuel.2023.128618.
Kolakoti, A., Tadros, M., Ambati, V.K., Gudlavalleti, V.N.S., 2023b. Optimization of biodiesel production, engine exhaust emissions, and vibration diagnosis using a combined approach of definitive screening design (DSD) and artificial neural network (ANN). Environ. Sci. Pollut. Control Ser. 30, 87260–87273. https://doi.org/10.1007/s11356-023-28619-1.
Kozina, A., Radica, G., Nižetić, S., 2020. Analysis of methods towards reduction of harmful pollutants from diesel engines. J. Clean. Prod. 262, 121105. https://doi.org/10.1016/j.jclepro.2020.121105.
Lafont-Torio, J., Martín, J.M., Salinas Fernández, J.A., Ribeiro-Soriano, D., 2023. Perceptions of progress toward achieving the sustainable development goals: insights from cooperative managers. Sustain. Technol. Entrepreneurship 3 (1), 100055. https://doi.org/10.1016/j.stae.2023.100055.
Leng, L., Li, T., Zhan, H., Rizwan, M., Zhang, W., Peng, H., Yang, Z., Li, H., 2023. Machine learning-aided prediction of nitrogen heterocycles in bio-oil from the pyrolysis of biomass. Energy 278, 127967. https://doi.org/10.1016/j.energy.2023.127967.
Su, W.-H., 2020. Advanced machine learning in point spectroscopy, RGB- and hyperspectral-imaging for automatic discriminations of crops and weeds: a review. Smart Cities 3, 767–792. https://doi.org/10.3390/smartcities3030039.
Tamilselvan, D., Sudhakar, T.D., 2024. Optimizing renewable energy utilization with high gain converters. Renew. Sustain. Energy Rev. 191, 114105. https://doi.org/10.1016/j.rser.2023.114105.
Torres-García, M., García-Martín, J.F., Jiménez-Espadafor Aguilar, F.J., Barbin, D.F., Álvarez-Mateos, P., 2020. Vegetable oils as renewable fuels for power plants based on low and medium speed diesel engines. J. Energy Inst. 93 (3), 953–961. https://doi.org/10.1016/j.joei.2019.08.006.
Yagmur, N., Dag, I., Temurtas, H., 2023. A new computer-aided diagnostic method for classifying anaemia disease: hybrid use of Tree Bagger and metaheuristics. Expert Syst. 41 (8), e13528. https://doi.org/10.1111/exsy.13528.
Yang, K., Wu, K., Zhang, H., 2022. Machine learning prediction of the yield and oxygen content of bio-oil via biomass characteristics and pyrolysis conditions. Energy 254, 124320. https://doi.org/10.1016/j.energy.2022.124320.
Yang, X., Wu, T., Huang, F., 2022. Reversible data hiding in JPEG images based on coefficient-first selection. Signal Process. 200, 108639. https://doi.org/10.1016/j.sigpro.2022.108639.
Zhang, J., Su, R., Fu, Q., Ren, W., Heide, F., Nie, Y., 2022. A survey on computational spectral reconstruction methods from RGB to hyperspectral imaging. Sci. Rep. 12, 11905. https://doi.org/10.1038/s41598-022-16223-1.
Zhang, T., Cao, D., Feng, X., Zhu, J., Lu, X., Mu, L., Qian, H., 2022. Machine learning prediction of bio-oil characteristics quantitatively relating to biomass compositions and pyrolysis conditions. Fuel 312, 122812. https://doi.org/10.1016/j.fuel.2021.122812.