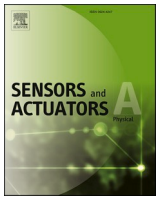
Sensors & Actuators: A. Physical 370 (2024) 115265

Contents lists available at ScienceDirect

Sensors and Actuators: A. Physical

journal homepage: www.journals.elsevier.com/sensors-and-actuators-a-physical



Size, material type, and concentration estimation for micro-particles in liquid samples

Sinan Genc, Talha Erdem *, Kutay Icoz *

Abdullah Gül University, Department of Electrical Electronics Engineering, Kayseri, Turkey

ARTICLE INFO

Keywords:
Optical Sensors
Micro plastic detection
Random forest
Mie scattering

ABSTRACT

The on-site examination and characterization of microparticles are becoming crucial due to the significant rise in plastic pollution in natural resources. Hence, identifying the specific microplastic composition and quantity would enable the implementation of preventive measures. This paper presents a cost-effective setup that utilizes the Random Forest algorithm to detect the size and refractive index of micro particles, hence facilitating the identification of the material type. The system utilizes the scattering patterns of laser light from the dispersion of microparticles, namely within the concentration range of 0.05 fM to 3.00 fM. The refractive indices and particle sizes of melamine (Me8) spheres with a size of 8 μm, as well as polystyrene (PS8) spheres with a size of 8 μm and (PS10) 10 μm, were estimated using the Random Forest algorithm and recorded scattering patterns. The proposed method may deliver findings with an average deviation of 0.23 μm for particle size and 0.015 for particle refractive index. The statistical analysis indicated that there was no notable disparity between the experimental findings and the predictions derived from the machine learning system. The existing configuration can be readily converted into a point-of-use system that can be employed on-site for the purpose of monitoring and identifying microplastic contamination.

1. Introduction

The thoughtless and widespread use of microplastics in our daily lives has led to the emergence of a global problem of microplastic pollution. The quantity of plastic waste in oceans and lakes is seeing a substantial rise as a result of the notable consumption of plastics [1–3]. The microparticles present in water mostly originate from two sources: fragmented plastic detritus and plastic microbeads often employed in various products. Upon pollution of the water, microplastics amass among fish and other marine organisms [4–6]. Plastics derived from automotive tires, industrial activities, and other sources significantly increase the quantity of these contaminants [7–9]. Recent scientific research has demonstrated the presence of these minuscule particles in the human body, blood, milk, and meat of farm animals. Additionally, they have been found in many food products such as honey, sugar, salt, and seafood [10–13]. A separate investigation revealed that minute submerged plastics of micro/nanoscale ascend to the water's surface and become intermingled with the air due to the process of evaporation [14]. Precipitation transports microplastics from the atmosphere back to the ground, leading to unavoidable contact and inhalation of these particles. The risk of inhaling microplastics escalates as their size decreases [15]. In order to successfully address the global issue of microplastic pollution, it is essential to understand its sources, possible impacts, and develop effective methods for detection and analysis. The widespread and careless utilization of plastics is resulting in a substantial surge in the quantity of plastic garbage in oceans and lakes. The consequences of this extend beyond the environment, as plastic waste is now infiltrating the human body through food consumption. Consequently, research endeavors have redirected towards innovative techniques such as light scattering to detect and examine microplastics, facilitating a more comprehensive understanding of their dimensions, prevalence, and composition. By employing techniques such as the Mie theory and utilizing advanced algorithms like Random Forest, these methods enable the implementation of preventive measures against the infiltration of microplastics into the human body, highlighting the urgent necessity to address this pervasive environmental issue.

Light scattering techniques can be employed to ascertain the size, composition, and concentration of different materials through diverse applications [16], [17]. To detect microplastics, one can measure and analyze the angle, distribution, and intensity of the scattered light.

* Corresponding authors.
E-mail addresses: erdem.talha@agu.edu.tr (T. Erdem), kutay.icoz@agu.edu.tr (K. Icoz).

https://doi.org/10.1016/j.sna.2024.115265
Received 14 August 2023; Received in revised form 11 December 2023; Accepted 7 March 2024
Available online 9 March 2024
0924-4247/© 2024 Elsevier B.V. All rights reserved.

Nevertheless, the Mie theory is constrained by the scattering of individual particles [18–20], and the inclusion of intricate equations such as Riccati-Bessel and Hankel leads to computational intricacy. An alternate approach involves implementing these equations on a solver and performing numerical calculations. By employing the numerical solution, one can acquire the scattering distribution of a particle by inputting the essential parameters, including particle dimensions, incident wavelength, and refractive indices for both the medium and the microparticle. In order to characterize the microplastics found in water-based solutions, we examine the pattern of scattered light and extract information about the size, refractive index, and concentration of the particles. The research we conducted is dependent on the angles of the interference rings, which were derived by theoretical calculations utilizing Mie scattering. Initially, we acquire bright interference ring angles which are dependent on particle size, refractive index, and wavelength of incident light. Afterwards, we utilize the interference ring angles that were obtained by numerical methods to train a Random Forest algorithm. Subsequently, the Random Forest technique is used to analyze the particle size and refractive index of scattering patterns obtained from experimental measurements of microspheres.

Transitioning from the analysis of microplastics and their characteristics to the broader landscape of particle detection, various techniques have been explored for their efficiency and portability. Literary research focuses on developing cost-efficient and portable instruments for particle detection. The particles and their properties were examined using various techniques such as Raman scattering, flow cytometry, fluorescence spectroscopy, and Fourier-transform infrared scattering [6], [16], [21–23]. Furthermore, to address the problem of microplastics pollution, various tiny and portable devices employing diverse methodologies have been created [24–28]. In addition, to develop the status of the technology, there are studies investigating the effects of particle shape and particles' distribution in the samples [29–31]. Having monodispersed particles in the samples is vital for accurately analyzing the impact of particle size. Developing a technique capable of identifying the dimensions, composition, and abundance of particles encountered in practical scenarios is crucial. In light of the escalating prevalence of microplastics and other undesirable contaminants, there will be a growing demand for solutions that are practical, cost-effective, portable, and robust.

By integrating machine learning into the detection process, it becomes feasible to analyze the vast quantity of photos and examine the dispersion of several particles. The Random Forest algorithm yields rapid and precise outcomes for scattering computations. The Random Forest method incorporates numerous trees, making it well-suited for our specific circumstances. Every tree incorporates judgments and thresholds that are considered during the classification of the data [32–34]. The collection of trees within the algorithm constitutes the forest, and the criteria for making decisions fluctuate randomly following each iteration. In this approach, testing is characterized as a stochastic process, resulting in varying outcomes with each repetition. The accuracy of the method is quantified by the standard deviation of the mean of all the outputs obtained during each iteration.

This study presents a cost-effective technique that enables the characterization of light scattering patterns by different microplastic samples in water, while considering the angles of interference rings. The studies utilized melamine particles with a diameter of 8 micrometers, as well as polystyrene particles with diameters of 8 micrometers and 10 micrometers. These particles were subjected to red laser light with a wavelength of 656.3 nanometers, green laser light with a wavelength of 514.9 nanometers, and blue laser light with a wavelength of 403.8 nanometers. The integrated system of our random forest is capable of delivering findings with an average deviation of 0.23 μm for particle size and 0.015 for particle refractive index, when comparing theory and predictions. Hence, we believe that the identification and characterization of the samples using a cost-efficient setup would provide a novel approach for the advancement of affordable and widely adopted technology to address the issue of global microplastic contamination.

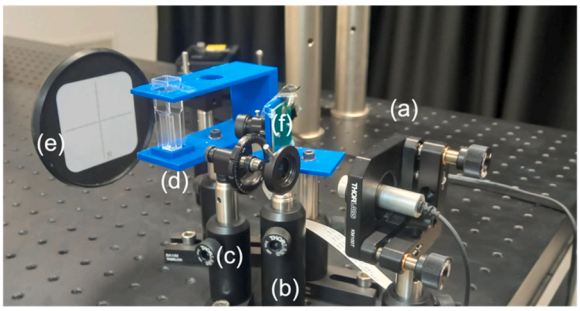


Fig. 1. Experimental setup, (a) light source, (b) iris, (c) neutral density filter, (d) cuvette holder and cuvette, (e) graded white screen, and (f) camera module.

2. Materials & methods

2.1. Sample preparation

It is feasible to examine the influence of material type (refractive index of particles) and particle size by utilizing commercially accessible microparticles composed of two distinct materials, melamine (Me) and polystyrene (PS), with diameters of 8 μm (PS8) and 10 μm (PS10), respectively (SIGMA-ALDRICH 95523, 84192, 72986). To investigate the correlation between particle quantity and scattering characteristics, samples at different concentrations were prepared.

Pipettes were utilized to fill the vials with the required quantity of stock solutions and ultra-pure water. The experiments included three types of materials, Me8, PS8, and PS10. The concentration ranged from 0.05 fM to 3.00 fM, or from 3×10^4 particles/mL to 1.8×10^6 particles/mL.

2.2. Experimental setup

A laser beam was directed towards the sample cuvette, which was at room temperature, after passing through an iris and a neutral density filter. Both the iris and the neutral density filter were employed for regulating the power and width of the incident beam. A graded screen was placed behind the micro-cuvette, which was fixed on a cuvette holder that was fabricated via 3D printing, as depicted in Fig. 1. A total of 200 scattering photos were captured for each sample under dark conditions using the Raspberry Pi Camera Module and a Raspberry Pi equipped with 4 GB of RAM. For the experiments, three lasers were utilized, each generating light at distinct wavelengths of 405 nm, 520 nm, and 650 nm from THORLABS (CPS405, CPS520, and CPS650F, respectively). Nevertheless, the data sheets provided by THORLABS indicate that the emission wavelengths of the lasers are at 403.8 nm, 514.9 nm, and 656.3 nm. To avoid potential confusions that other researchers, who want to employ our results or methods, may experience, we have opted to utilize these measured values reported by THORLABS.

2.3. Image processing and machine learning

The 200 captured images of each sample were transmitted to a computer for additional processing. Initially, a mean image was acquired to diminish the level of noise, as depicted in Fig. 2(a). The image exhibits symmetry and was resized to one-fourth of its original dimensions. Additionally, it was transformed into grayscale to optimize computing efficiency (Fig. 2(b-c)). To collect data on angular scattering, azimuthal angles ranging from 5 to 90° were selected for integration, with a step size of 1°, as shown in Fig. 2(d). Azimuthal angles pertain to the angles on the plane parallel to the screen (i.e., positive x-axis shown on the screen correspond to 0° whereas positive y-axis correspond to 90° (Fig. 1(e)), whereas scattering angles are the angles with respect to the

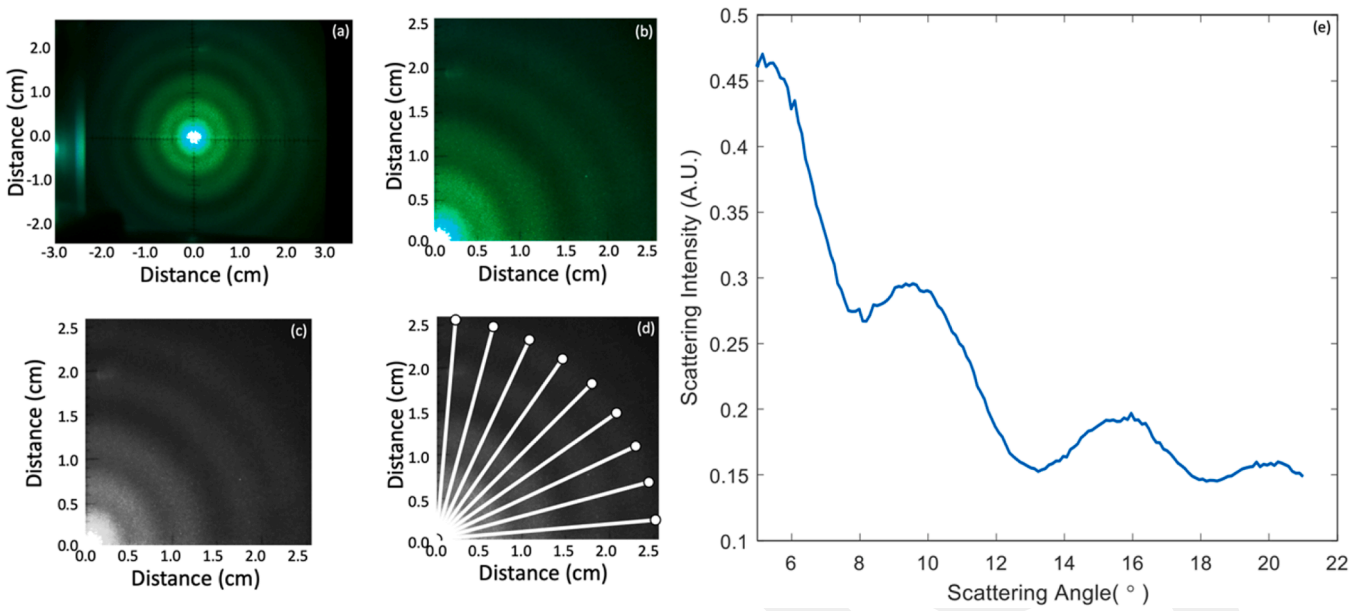


Fig. 2. (a) Raw scattering image of Me8 particles excited by green laser (b) cropped image (c) gray-scale image (d) data lines on scattering image (e) average scattering behavior of 86 lines (5°:1°:90°) on 1.50 fM 8 μm Me particles.

axis aligned with the light source (i.e., the angle that the scattered light makes with the optical axis). In Fig. 2(d), the 86 azimuthal lines were depicted using 10° intervals to enhance clarity of the illustration. Ultimately, the angular scattering data of the sample was analyzed. Fig. 2(e) displays the results, which exhibit constructive interference rings appearing as peaks in relation to the scattering angle. To enhance the accuracy of peak detection, the data points around the peak angles were used to fit a second-degree polynomial, resulting in more stable peak angle values, and reducing uncertainty. The scattering angle varies from 0° to 21°, which was selected due the optimal distance between the cuvette and the screen so that intensity variations and scattering rings could be reliable imaged by our camera. Modifying the distance would result in different scattering ranges. When the screen is placed too close to the sample, the central area appears too bright. This results in improper identification of the scattering at larger angles since they appear much weaker on the screen compared to the central part of the scattered pattern. Furthermore, exploring scattering at broader angles would result in a higher computing burden without providing any significant improvement to the quality of the data. On the other hand, placing the screen far from the sample results in imaging a lower number of scattering rings, which leads to a loss of crucial information related to particle characteristics. Since we found out that the first four scattering rings were enough to reach conclusions related to particle characteristics, we have chosen this scattering range in this study.

Given the inherent numerous scatterings of the spheres in liquid, the Random Forest algorithm offers the opportunity to categorize the measured scattering data, which would not be a simple theoretical calculation. Random Forest was utilized in this investigation because of its straightforward design and repeatable outcomes. Using Mie scattering equations existing in the literature [18], [20], [35], we calculated theoretical scattering angles by a numerical solver. The correspondence between experimental findings and numerical solutions were investigated and the results are given in Supplementary Information, Table S1 and Table S2. The highly matched results made it possible to apply Random Forest conveniently for creation of a dataset. This numerical solution methodology produced the data set, which included the wavelength of incident light, particle size, material type, and bright ring angles that were used to run the algorithm. A test set of data was created from 20% of all measurements. The remaining part was used to train the Random Forest algorithm.

The inputs for classifying the material types and particle sizes were incident wavelength and bright ring angles only. The concentration and scattering intensities were not required at this stage because it was observed during the image analysis that the peaks were at the same angle for a given material type, concentration, and size while the intensity of the scattered light increases with increasing concentration. In addition, particle size was not used as an input parameter to predict particle refractive index or vice versa.

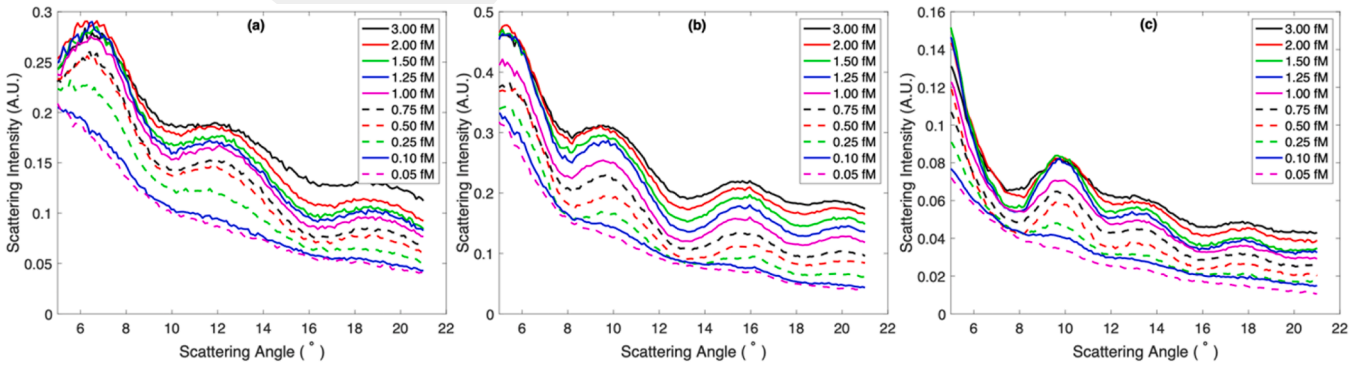


Fig. 3. Angular scattering of Me8 particles for concentrations from 0.05–3.00 fM by (a) 656.3 nm, (b) 514.9 nm and (c) 403.8 nm.

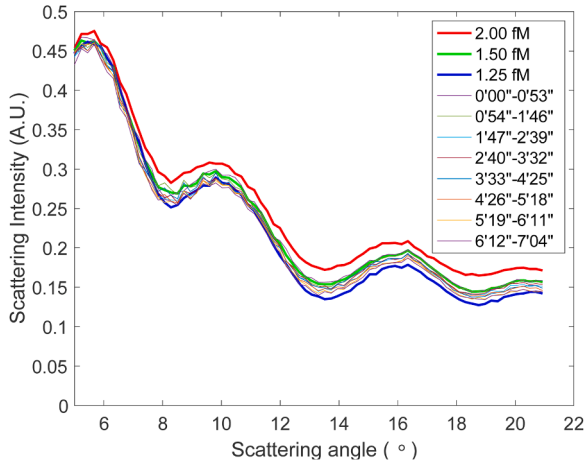


Fig. 4. Scattering data of 2.00 fM, 1.25 fM and averages of each consecutive 100 images belong to the 1.50 fM Me8 particles for a total of 800 images.

Univariate normality is assessed by examining the skewness and kurtosis values. The skewness value quantifies the degree of symmetry in a distribution, while the kurtosis value quantifies the degree of peakedness [36]. The allowable range for skewness and kurtosis is between −2 and +2 [37]. Furthermore, when the skewness and kurtosis values are close to zero, it indicates that the distribution is normal [38]. Levene's test was used to verify the homogeneity of the variances. Verifying and guaranteeing these two assumptions is crucial as the subsequent statistical tests are parametric. The significance of the statistical results was evaluated at a significance level of $p<0.05$. An independent sample t-test was conducted to assess the consistency of the projected outcomes in both the testing and experimental portions of the dataset.

3. Results and discussion

Consistent with expectations, the intensity of the scattered light increases in tandem with the number of particles (Fig. 3), particularly from a concentration of 0.25 fM onwards (Fig. S1 and S2 for PS8 and PS10). It is evident that variations in concentration have no substantial impact on the scattering peak angles, provided that the particle's refractive index, size, and the incident light's wavelength remain constant.

The examination of particle sedimentation at the bottom of the cuvette was essential for accurate measurements, despite the fact that capturing 200 images of each sample is a brief process, taking approximately 1.5 minutes. Hence, we conducted an investigation to examine the impact of sedimentation on our measurements. We obtained 800 photographs, which is four times more than the typical number of images, of 8 μm Melamine particles that were triggered by a green laser. We then organized these images into groups of 100. According to the data shown in Fig. 4, the scattering intensity of the final set is greater than that of the nearest lower concentration, which is 1.25 fM. Thus, we may infer that the length of our imaging procedure and the settling of particles in the samples do not affect our approach.

The calibration curves for Me8 particles at incoming light wavelengths of 656.3 nm, 514.9 nm, and 403.8 nm are displayed in Fig. 5. These curves enable us to accurately ascertain the concentration of a sample. Additional calibration curves for PS8 and PS10 may be found in the Supplementary Information, namely in Figures S3 and S4, respectively. We generated three distinct sample sets consisting of identical types and concentrations of substances and carried out experiments at three different time points. The integration of all pixel values in a picture

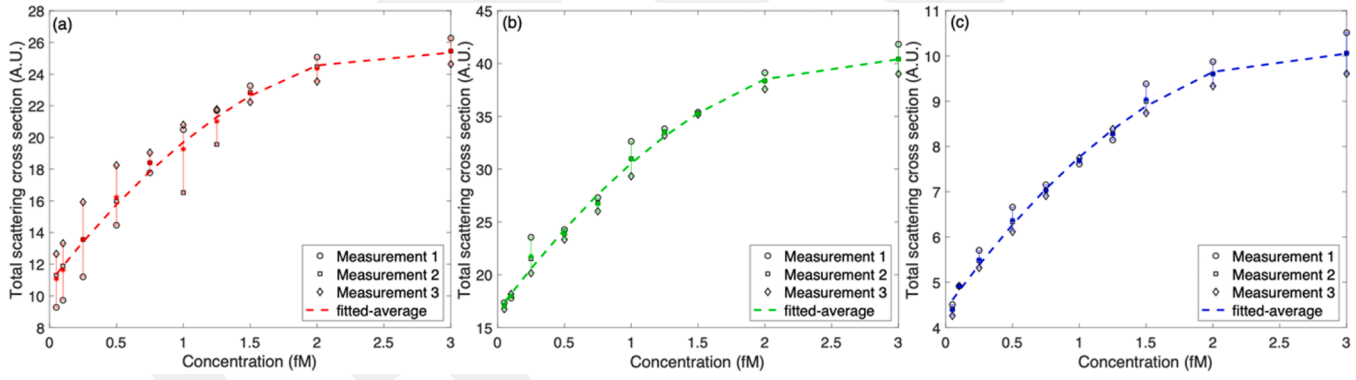


Fig. 5. Total scattering cross section and calibration curves of Me8 particles for three different sample sets and measurements by employing lasers emitting at (a) 656.3 nm, (b) 514.9 nm and (c) 403.8 nm.

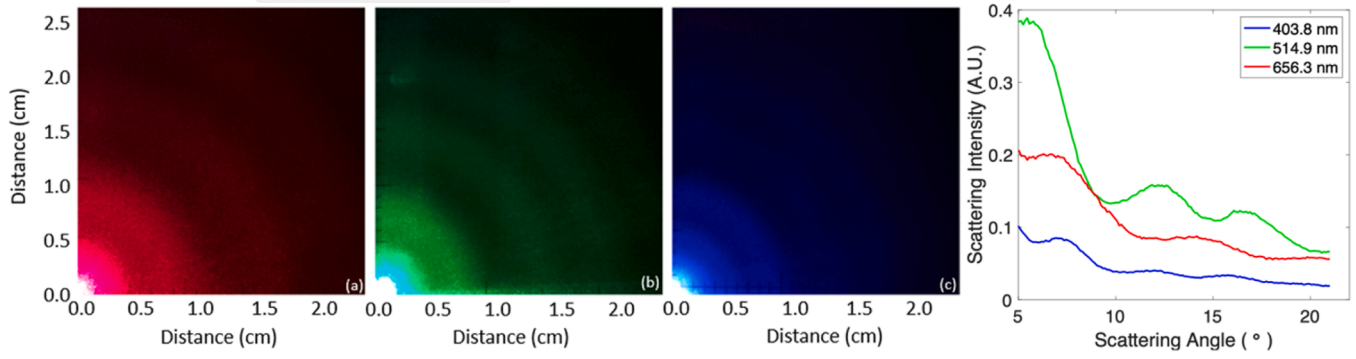


Fig. 6. Right-upper quartile of calculated scattering patterns of 8 μm polystyrene (PS8) particles by illuminating the samples with lasers emitting at (a) 656.3 nm, (b) 514.9 nm and (c) 403.8 nm. (d) shows the integrated scattering data as a function of angle.

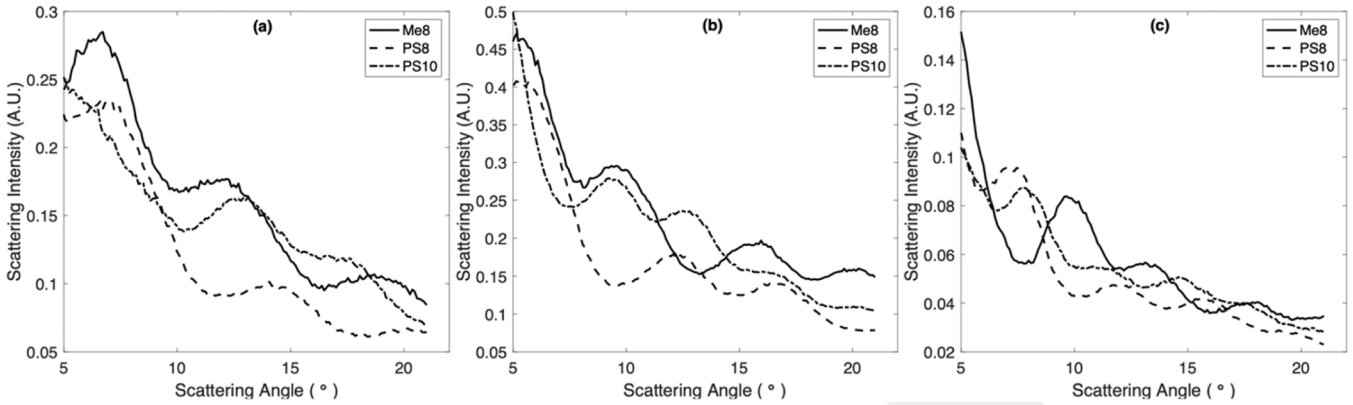


Fig. 7. Scattering data of 1.50 fM Me8, PS8 and PS10 particles by lasers emitting at (a) 656.3 nm, (b) 514.9 nm and (c) 403.8 nm.

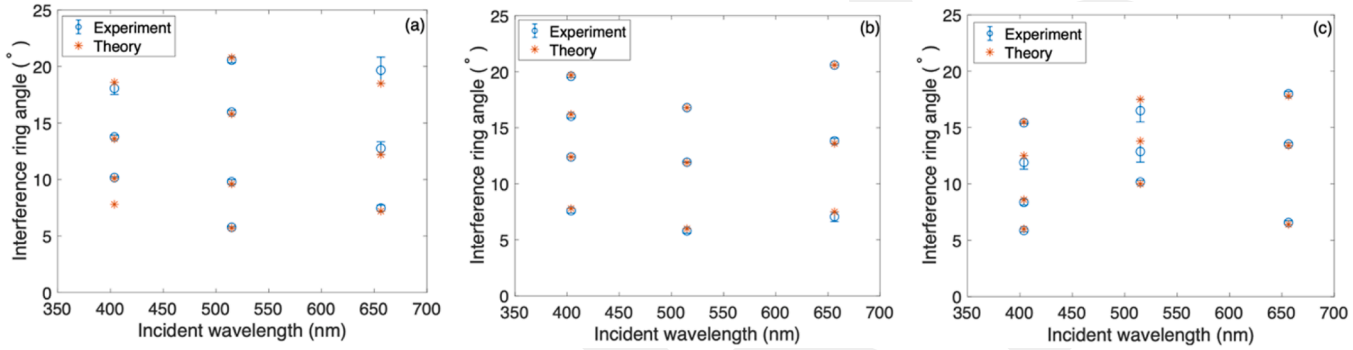


Fig. 8. Comparison of the first four peak angles of experimentally obtained scattering patterns with the patterns obtained using Mie theory for (a) Me8, (b) PS8, and (c) PS10 particles.

yields the cumulative scattered light intensity. We computed the integrated scattering intensities by analyzing the scattering patterns of three distinct samples for every concentration. Calibration curves were generated by applying a second-degree polynomial regression to the average concentrations of particles, sizes, and incident wavelengths.

As mentioned in the section on Image Processing and Machine Learning, the Random Forest algorithm generates random decision trees based on input data and ultimately determines the output value. In order to accomplish it, there must be a distinction between the input sets. It is essential in our study that the first four bright ring angles of a sample must be distinct. In order to assess the potential use of interference ring angles as input parameters for the Random Forest algorithm, we compared the fluctuations in incidence wavelength, particle size, and refractive index of the particle. By observing notable variations in scattering patterns, we are able to discern the dimensions, refractive index, and abundance of particles.

Three unique incident wavelengths were employed to investigate the scattering patterns of the same material (PS8 at a concentration of 1.00 fM) in order to analyze the findings related to the bright interference rings. The peaks generated by red, green, and blue lasers exhibited significant dissimilarities to each other, as depicted in Fig. 6, while using the identical material and concentration. Therefore, we have determined that varying incident wavelengths lead to distinct scattering patterns that rely on the angle. Moreover, the unique ring angles observed at the same concentration but with different incident light wavelengths facilitate the determination of particle size and wavelength-dependent refractive index. This, in turn, enables the identification of the particle type. Fig. 7 demonstrates that samples with a concentration of 1.50 fM exhibit data with different interference ring angles. Utilizing this information would be suitable for employing the Random Forest technique to aid in the categorization of the samples. It is evident that various materials exhibit distinct Mie resonances when exposed to light of the same wavelength.

We examined the potential overlap between the interference ring angles calculated using Mie theory and the corresponding experimental results (as shown in Fig. 8 and Table S1 (Supplementary Table 1)). Initially, the average peak angles of ten concentrations were evaluated for each combination of particles and light sources. The average of the first four peak angle values (1st, 2nd, 3rd, and 4th) was determined and compared to the theoretical values. This was done by calculating the mean absolute percentage error (MAPE) and computing the standard deviations (STD) based on the experimental measurements of three independent sample sets. Furthermore, the root mean squared error (RMSE) was computed by comparing the experimental results with the theoretical values. The RMSE has a maximum error range of 0.08–1.01, while the MAPE has a maximum error range of 0–6.76%. The computed errors and standard deviations are included in the Supplementary Information (Table S2 (Supplementary Table 2)). These findings indicate that incorporating the Random Forest algorithm into a cost-effective, portable device should yield satisfactory results in accurately predicting tiny discrepancies between real and anticipated outcomes. Although there are lesser variations in shorter wavelengths for 8 μm particles, the experimental findings for 10 μm PS particles align most closely with the theory at a wavelength of 656.3 nm. Furthermore, the largest difference of bright ring angles between the average of the experimental data and the theoretical calculations were obtained for Me particles irradiated with the red laser. This error is found to be 1.17°, as shown in the "Theory" columns of Table S1 and the "AVEG" columns of Table S2. In addition, the agreement between theory and experimentation is more accurate at smaller scattering angles compared to larger scattering angles. Hence, it is reasonable to conclude that the theoretical calculations are applicable to the experimental configuration.

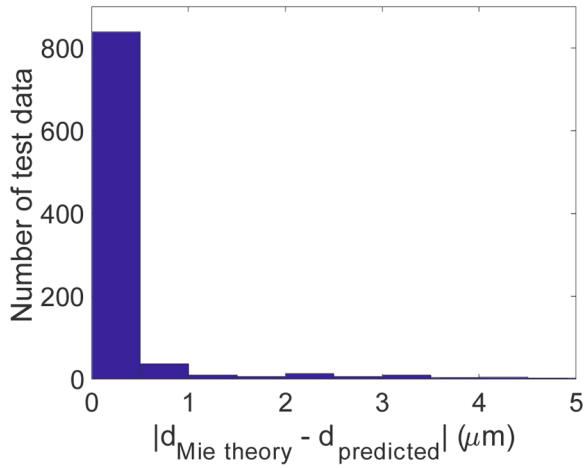


Fig. 9. The difference between the Mie theory and the predicted results for the test data by the Random Forest algorithm.

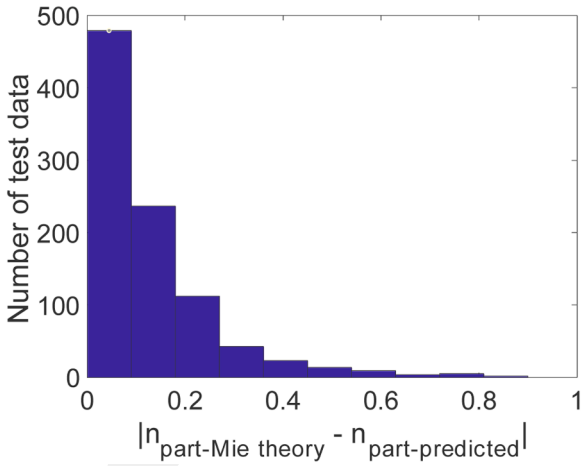


Fig. 10. The difference between the Mie theory and the predicted results for the test data by the Random Forest algorithm.

Although it has been demonstrated that microparticles exhibit distinct Mie resonances, characterized by varying bright interference ring angles and intensities, conducting experiments with numerous materials at different concentrations and sizes would necessitate a substantial amount of laboratory work to generate a dataset solely based on experimental findings. Alternatively, a dataset was generated by theoretical calculations that ensure conformity with experimental findings, and this dataset will be employed in the machine learning method. A dataset was built using nested for loops to account for particle size, material type, and incidence wavelength. The incidence wavelength ranges in our laboratory were 403.8 nm, 514.9 nm, and 656.3 nm. The particle refractive index varied from 1.30 to 2.20 in increments of 0.01. The particle diameter ranged from 5 μm to 13 μm in increments of 0.5 μm. We developed two distinct models for the purpose of determining particle size and refractive index. The only parameters known for those two models were the incident wavelength, the first four angles of the interference rings, and the refractive index of the medium. We did not include size information when searching for the refractive index of the particle, nor did we consider the refractive index when seeking its size.

For the training of both models, a random selection of 80% of the dataset was used. The model was provided with the incoming wavelength, the refractive index of the medium which is water in this study, and the angles of the first four brilliant rings, ranging from 5 to 21 degrees. The results provided data on the dimensions and refractive index of the particle. The Random Forest algorithm was implemented using 20 iterations, each consisting of 100 trees, using a methodology similar to that described in the literature [32], [33].

In order to assess the suitability of independent variable t-tests, it is necessary for the results to follow a normal distribution. The results indicate that the skewness and kurtosis values for the real data are 0.05 and −1.191, respectively, which fall within acceptable ranges. The test data exhibits skewness and kurtosis values of 0.026 and −1.144, respectively, which are within the permitted range for conducting an independent sample t-test. The significant value (0.741) of Levene's test for equality of variances indicates that the data satisfies the condition of homogeneity of variances. Based on the assumption of equal variances between the real and predicted data for Model 1, the result of the independent sample T-test within a 95% confidence range indicates that there is no statistically significant difference between the mean values of the real and predicted data ($t=-0.033$, $df=1854$, $p>0.05$).

The particle size was estimated using the refractive index of the medium, the incident wavelength, and the first four interference ring angles. The essential parameters that can be extracted from a raw image captured during the measurement are as follows. The remaining portion of the dataset, which accounts for 20% of the library, was employed to evaluate the model's performance subsequent to training it using those inputs. The prediction of particle size exhibited a correlation of 0.95 with 20% of the library, as depicted in Fig. 9. After establishing the accuracy value of the model utilizing the library generated by theoretical results, we utilized experimental results as test data for the model. The particles Me8, PS8, and PS10 have been utilized to examine the properties of red, green, and blue light. The actual and projected particle sizes for such combinations are illustrated in Table 1. The T-Test results, with a value of 0.973, indicate that there is no statistically significant distinction between the actual and projected particle sizes. The results demonstrate the efficacy of our technology in detecting the size of particles.

The second model utilized identical input parameters, including the refractive index of the medium, incident wavelength, and the first four angles of interference rings, to forecast the refractive index of particles during training. The identical methods, with a ratio of 80% for training data and 20% for testing data, were executed. The performance of the model in estimating the refractive index of particles is presented in Fig. 10. The prediction of particle size exhibited a correlation of 0.748 with 20% of the library, as depicted in Fig. 10. Providing varying sets of angles for the same material at different sizes poses a challenge for the model's learning process. For example, a material exhibits varying peak angles at different sizes, while maintaining a consistent refractive index. The training procedure is limited to using only incoming wavelength and peak angle values, without considering particle size. This is due to the fact that multiple sets of input parameters can provide the same output, making it difficult to accurately predict the forces involved. Thus, the distribution exhibits greater inaccuracy in comparison to the

Table 1
The comparison of actual and predicted results for particle size by the Random Forest algorithm.

	Red (656.3 nm)			Green (514.9 nm)			Blue (403.8 nm)		
	Me8	PS8	PS10	Me8	PS8	PS10	Me8	PS8	PS10
Actual	8 μm	8 μm	10 μm	8 μm	8 μm	10 μm	8 μm	8 μm	10 μm
Predicted	8 μm	6.5 μm	10 μm	8 μm	8 μm	10 μm	8 μm	8 μm	9.5 μm

Table 2
The comparison of the actual and the predicted results for refractive index of the particle by the Random Forest algorithm.

	Red (656.3 nm)			Green (514.9 nm)			Blue (403.8 nm)		
	Me8	PS8	PS10	Me8	PS8	PS10	Me8	PS8	PS10
Actual	1.79	1.589	1.589	1.89	1.601	1.601	1.96	1.617	1.617
Predicted	1.79	1.72	1.61	1.89	1.60	1.60	1.96	1.56	1.66

particle size.

Model 2 exhibited suitable skewness and kurtosis values, enabling the application of an independent sample t-test for this model. The significant value (0.711) of Levene's test for equality of variances indicates that the data satisfies the requirement of homogeneity of variances. Based on the assumption of equal variance between the real and predicted data for Model 2, the result of the independent sample T-test within a 95% confidence range indicates that there is no statistically significant difference between the mean values of the real and predicted data ($t=-0.202$, $df= 1854$, $p>0.05$).

While the discrepancy between laboratory and estimated findings is greater for the distribution of error compared to the particle size results, the second model nevertheless demonstrated satisfactory performance, taking into account the variations outlined in Table 2. Furthermore, the T-Test results (0.840) indicate that there is no statistically significant distinction between the actual and expected refractive index of the particles, similar to the particle size. Hence, the refractive indices of the materials can be accurately calculated by analyzing the scattering images captured using the suggested arrangement with a minimal margin of error. The system is capable of differentiating between two spheres of equal size by analyzing their refractive indices and scattering characteristics at specific wavelengths.

Our cost-effective experimental system is designed to detect certain particles that are recognized by the machine learning model. In our prior study, it was demonstrated that there is no substantial disparity between the outputs of the experimental system and the findings of numerical calculations [39]. Consequently, after the model has been trained using a diverse dataset, it will possess the capability to make separate predictions for particle size and refractive index. To mitigate the impact of diversity in particle size, shape, or refractive index, our method needs to be further improved by linking the scattering pattern with the effects of these variations in a sample. It may also be necessary to add supplementary procedures prior to analysis such as inclusion of filtering units with various pore sizes. Furthermore, the utilization of various angles to collect scattering data which would have differences due to asymmetries in the particle shapes and integration of this scattering information with image processing techniques have the potential for enabling particle shape assessments.

4. Conclusions

The issue of microplastic pollution is rapidly expanding and necessitates urgent measures to restrict and monitor the release of microplastics into our natural resources. The scattering theory is used for the purpose of detecting, identifying, and classifying particles present in liquids. Regrettably, performing theoretical computations for particle scattering manually poses a significant challenge. Furthermore, the process is impeded by the existence of multiple particles and the phenomenon of light scattering occurring among these particles.

This work involved capturing the scattering patterns of particles with varied sizes and compositions using an inexpensive apparatus. We conducted experiments using three distinct wavelengths of input light. Initially, we established a direct relationship between the Mie theory and the empirical evidence. The experimental results we obtained were consistent with the results predicted by the theoretical equations implemented in a numerical solver. The largest disparity seen between experimental and theoretical outcomes for the angle of a bright ring was 1 degree, as demonstrated in the "Theory" columns of Table S1 and the "AVEG" columns of Table S2. Subsequently, we generated a compilation of scattering patterns by employing Mie theory. Alternatively, doing experiments with numerous materials would result in significant expenditures of time, resources, and computational effort. Furthermore, there remains a constraint imposed by the unused materials for data collection. Nevertheless, the utilization of theoretical findings that align with experimental outcomes has led to the emergence of prospects for additional advancements. Consequently, we proceeded to create a machine learning model to estimate the dimensions of particles and their refractive index. We have created a predictive model for estimating particle size and refractive index. This model is able to generate certain scattering patterns. Additionally, we have devised a technique to extract relevant data from these scattering patterns. This method aims to address computational challenges and minimize the time required for calculations. Through our analysis of the scattering images, we established that the interference ring angles exhibited unique values. The algorithm might utilize varying bright ring angles corresponding to different materials and particle sizes as its inputs. Consequently, they might be utilized as inputs for the machine learning model. Following the examination of our laboratory findings and the acquisition of precise predictions, we have ultimately achieved a high level of reliability. The Random Forest technique accurately obtained 7/9th of the particle size in model 1. The largest discrepancy observed was a deviation of 1.5 micrometers specifically for PS8 when exposed to red light. The maximum inaccuracy for the refractive index of particles in model 2 is 0.13. The outcomes obtained with Me8 using green light are nearly indistinguishable from the actual values. Therefore, it has been demonstrated that this configuration is capable of discerning the material composition and dimensions of two spheres of equal size.

Considering the setup's affordable price and ease of transport, we believe that this work will introduce fresh possibilities for particle detection and investigation, with significant ramifications for assessing water quality.

CRediT authorship contribution statement

Sinan Genc: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Talha Erdem:** Conceptualization, Data curation, Funding acquisition, Methodology, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Kutay Icoz:** Conceptualization, Funding acquisition, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

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.

Data Availability

Data will be made available on request.

Acknowledgments

TE acknowledges BAGEP. SG acknowledges Dr. Serhat Ozer for his feedback about statistical analysis.

Declaration of AI use

We have not used AI-assisted technologies in creating this article.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.sna.2024.115265.

References

[1] E.J. Carpenter, K.L. Smith, Plastics on the Sargasso sea surface, Science (1979) vol. 175 (4027) (1972) 1240–1241, https://doi.org/10.1126/SCIENCE.175.4027.1240.
[2] E.J. Carpenter, S.J. Anderson, G.R. Harvey, H.P. Miklas, B.B. Peck, Polystyrene spherules in coastal waters, Science (1979) vol. 178 (4062) (1972) 749–750, https://doi.org/10.1126/SCIENCE.178.4062.749.
[3] A. Ashkin, Acceleration and trapping of particles by radiation pressure, Phys. Rev. Lett. vol. 24 (4) (1970) 156, https://doi.org/10.1103/PhysRevLett.24.156.
[4] K.D. Cox, G.A. Covernton, H.L. Davies, J.F. Dower, F. Juanes, S.E. Dudas, Human consumption of microplastics, Environ. Sci. Technol. vol. 53 (12) (2019) 7068–7074, https://doi.org/10.1021/ACS.EST.9B01517.
[5] A.Dick Vethaak, J. Legler, Microplastics and human health, Science (1979) vol. 371 (6530) (2021) 672–674, https://doi.org/10.1126/SCIENCE.ABE5041.
[6] D. Schymanski, C. Goldbeck, H.U. Humpf, P. Fürst, Analysis of microplastics in water by micro-Raman spectroscopy: release of plastic particles from different packaging into mineral water, Water Res vol. 129 (2018) 154–162, https://doi.org/10.1016/J.WATRES.2017.11.011.
[7] A. Isobe, S. Iwasaki, K. Uchida, T. Tokai, Abundance of non-conservative microplastics in the upper ocean from 1957 to 2066, Nat. Commun. vol. 10 (1) (2019) 1–13, https://doi.org/10.1038/s41467-019-08316-9.
[8] I.E. Napper, A. Bakir, S.J. Rowland, R.C. Thompson, Characterisation, quantity and sorptive properties of microplastics extracted from cosmetics, Mar. Pollut. Bull. vol. 99 (1–2) (2015) 178–185, https://doi.org/10.1016/J.MARPOLBUL.2015.07.029.
[9] K.E. Peiponen, J. Räty, U. Ishaq, S. Pélisset, R. Ali, Outlook on optical identification of micro- and nanoplastics in aquatic environments, Chemosphere vol. 214 (2019) 424–429, https://doi.org/10.1016/J.CHEMOSPHERE.2018.09.111.
[10] Y. Su, et al., Author Correction: Steam disinfection releases micro(nano)plastics from silicone-rubber baby teats as examined by optical photothermal infrared microspectroscopy, Nat. Nanotechnol. vol. 17 (6) (2022), https://doi.org/10.1038/s41565-022-01155-8.
[11] Q. Zhang, et al., A review of microplastics in table salt, drinking water, and air: direct human exposure, Environ. Sci. Technol. vol. 54 (7) (2020) 3740–3751, https://doi.org/10.1021/ACS.EST.9B04535.
[12] L.M. Hernandez, E.G. Xu, H.C.E. Larsson, R. Tahara, V.B. Maisuria, N. Tufenkji, Plastic teabags release billions of microparticles and nanoparticles into tea, Environ. Sci. Technol. vol. 53 (21) (2019) 12300–12310, https://doi.org/10.1021/ACS.EST.9B02540.
[13] D. Li, et al., Microplastic release from the degradation of polypropylene feeding bottles during infant formula preparation, Nat. Food vol. 1 (11) (2020) 746–754, https://doi.org/10.1038/s43016-020-00171-y.
[14] S. Sridharan, M. Kumar, L. Singh, N.S. Bolan, M. Saha, Microplastics as an emerging source of particulate air pollution: a critical review, J. Hazard Mater. vol. 418 (2021) 126245, https://doi.org/10.1016/J.JHAZMAT.2021.126245.
[15] P. Orellano, J. Reynoso, N. Quaranta, A. Bardach, A. Ciapponi, Short-term exposure to particulate matter (PM10 and PM2.5), nitrogen dioxide (NO2), and ozone (O3) and all-cause and cause-specific mortality: systematic review and meta-analysis, Environ. Int vol. 142 (2020), https://doi.org/10.1016/J.ENVINT.2020.105876.
[16] N. Kaile, M. Lindivat, J. Elio, G. Thuestad, Q.G. Crowley, I.A. Hoell, Preliminary results from detection of microplastics in liquid samples using flow cytometry, Front Mar. Sci. vol. 7 (2020) 856, https://doi.org/10.3389/FMARS.2020.552688/BIBTEX.
[17] S. Veerasingam, et al., Contributions of fourier transform infrared spectroscopy in microplastic pollution research: a review, Crit. Rev. Environ. Sci. Technol. vol. 51 (22) (2020) 2681–2743, https://doi.org/10.1080/10643389.2020.1807450.
[18] H.C. van de Hulst, Light Scattering by Small Particles, Dover Publications, 1981.
[19] J.R. Frisvad, N.J. Christensen, H.W. Jensen, Predicting the apperance of materials using Lorenz-Mie theory, in: W. Hergert, T. Wriedt (Eds.), The Mie Theory, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, https://doi.org/10.1007/978-3-642-28738-1.
[20] C.F. Bohren, D.R. Huffman, Absorption and Scattering of Light by Small Particles, Wiley, 1983.
[21] Y. Chen, et al., Identification and quantification of microplastics using Fourier-transform infrared spectroscopy: Current status and future prospects, Curr. Opin. Environ. Sci. Health vol. 18 (2020) 14–19, https://doi.org/10.1016/J.COESH.2020.05.004.
[22] A. Scircle, J.V. Cizdziel, Detecting and quantifying microplastics in bottled water using fluorescence microscopy: a new experiment for instrumental analysis and environmental chemistry courses, J. Chem. Educ. vol. 97 (1) (2020) 234–238, https://doi.org/10.1021/ACS.JCHEMED.9B00593.
[23] L. Cabernard, L. Roscher, C. Lorenz, G. Gerdts, S. Primpke, Comparison of Raman and Fourier Transform Infrared Spectroscopy for the quantification of microplastics in the aquatic environment, Environ. Sci. Technol. vol. 52 (22) (2018) 13279–13288, https://doi.org/10.1021/ACS.EST.8B03438.
[24] A.H. Iri, et al., Optical detection of microplastics in water, Environ. Sci. Pollut. Res. vol. 28 (45) (2021) 63860–63866, https://doi.org/10.1007/S11356-021-12358-2.
[25] C.K. Chen, et al., A portable purification system for the rapid removal of microplastics from environmental samples, Chem. Eng. J. vol. 428 (2022) 132614, https://doi.org/10.1016/J.CEJ.2021.132614.
[26] X. Zhang, et al., Rapid monitoring approach for microplastics using portable pyrolysis-mass spectrometry, Anal. Chem. vol. 92 (6) (2020) 4656–4662, https://doi.org/10.1021/ACS.ANALCHEM.0C00300.
[27] B.O. Asamoah, B. Kanyathare, M. Roussey, K.E. Peiponen, A prototype of a portable optical sensor for the detection of transparent and translucent microplastics in freshwater, Chemosphere vol. 231 (2019) 161–167, https://doi.org/10.1016/J.CHEMOSPHERE.2019.05.114.
[28] R.L. Coppock, M. Cole, P.K. Lindeque, A.M. Queirós, T.S. Galloway, A small-scale, portable method for extracting microplastics from marine sediments, Environ. Pollut. vol. 230 (2017) 829–837, https://doi.org/10.1016/J.ENVPOL.2017.07.017.
[29] M. Lin, Z. Ming, C. Li, Aerosol Sauter mean diameter measurement based on the light scattering response of the combined particle volume-surface area, Opt. Express vol. 31 (3) (2023) 3490–3503, https://doi.org/10.1364/OE.477231.
[30] M. Lin, M. Zhu, C. Li, Y. Chen, In situ optical sensor for aerosol ovality and size, Sens Actuators A Phys. vol. 347 (2022) 113963, https://doi.org/10.1016/J.SNA.2022.113963.
[31] M. Lin, M. Zhu, X. Xiao, C. Li, J. Wu, Optical sensor for combustion aerosol particle size distribution measurement based on embedded chip with low-complexity Mie scattering algorithm, Opt. Laser Technol. vol. 158 (2023) 108791, https://doi.org/10.1016/J.OPTLASTEC.2022.108791.
[32] J.L. Speiser, M.E. Miller, J. Tooze, E. Ip, A comparison of random forest variable selection methods for classification prediction modeling, Expert Syst. Appl. vol. 134 (2019) 93–101, https://doi.org/10.1016/J.ESWA.2019.05.028.
[33] G. Biau, E. Scornet, A random forest guided tour, Test vol. 25 (2) (2016) 197–227, https://doi.org/10.1007/S11749-016-0481-7.
[34] M.Z. Joharestani, C. Cao, X. Ni, B. Bashir, S. Talebiesfandarani, PM2.5 prediction based on random forest, XGBoost, and deep learning using multisource remote sensing data, Atmosphere vol. 10 (7) (2019) 373, https://doi.org/10.3390/ATMOS10070373.
[35] D.W. Hahn, "Light Scattering Theory," 2009.
[36] B.G. Tabachnick, L.S. Fidell. Using Multivariate Statistics, 7th ed., Pearson Education, Inc, 2019.
[37] Darren George, P. Mallery, SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update, Allyn & Bacon, 2010.
[38] J.F. Jr. Hair, W.C. Black, B.J. Babin, and R.E. Anderson, *Multivariate Data Analysis* (7th Edition) by Joseph F. Hair, William C. Black, Barry J. Babin, Rolph E. Anderson, Pearson Education, Inc., 2010.
[39] S. Genc, K. Icoz, T. Erdem, Numerical analysis and experimental verification of optical scattering from microplastics, R. Soc. Open Sci. vol. 10 (8) (2023), https://doi.org/10.1098/RSOS.230586.

Dr. Sinan Genc is a research assistant at Abdullah Gül University. He received his BSc in electrical electronics engineering from Sakarya University in 2012, MSc and PhD degrees in electrical and computer engineering from Abdullah Gül University in 2016 and 2023, respectively. His Postdoctoral studies are going on at Middle East Technical University. His current research interests include optical sensors, particle detection, characterization, and optoelectronic systems.

Dr. Talha Erdem received his BS, MS, and PhD degrees all in Electrical and Electronics Engineering from Bilkent University in 2009, 2011, and 2016, respectively. After his PhD he was awarded the Newton International Fellowship by the Royal Socıety and moved to the University of Cambridge as a Newton International Fellow. Since April 2019 he has been working at Abdullah Gül University as an Assistant Professor and established the Smart Nanophotonics Research Group, His current research interests are the design of stable nano-emitters, DNA-driven self-assembly of colloidal nanoparticles, and their photonic applications.

Dr. Kutay Icoz received the Ph.D. degree from Purdue University Biomedical Engr. Department in 2010. He worked at Harvard Medical School and Massachusetts General Hospital Department of Neurosurgery as a postdoctoral research fellow and at Intel Corporation Assembly & Test Technology Development Division as a senior engineer. He currently works as a faculty member of Electrical-Electronics Engineering at Abdullah Gül University. His research focuses on novel applications of micro/nano technology on biology and medicine, biosensors, point of care devices and wearable biomedical devices.