20

# Optical identification of hemolysis and lipemia in blood serum samples: computer vision and diffuse reflectance spectroscopy

© G.M. Denisenko[1,2], R.R. Fitagdinov[3,4,5], B.P. Yakimov[2,6,7], A.A. Biryukov[7,8], Yu.A. Shitova[5], E.N. Keruntu[7], A.S. Shkoda[7], E.A. Shirshin[2,6]

[1] Institute for Regenerative Medicine, Sechenov First Moscow State Medical University, 119991 Moscow, Russia
[2] Laboratory of Clinical Biophotonics, Biomedical Science & Technology Park, Sechenov First Moscow State Medical University, 119991 Moscow, Russia
[3] Moscow Institute of Physics and Technology (State University), 141701 Dolgoprudnyi, Moscow oblast, Russia
[4] Institute for Nuclear Research, Russian Academy of Sciences, 117312 Moscow, Troitsk, Russia
[5] JSC „Medeum“, 127566 Moscow, Russia
[6] Moscow State University, Faculty of Physics, 119234 Moscow, Russia
[7] Vorohobov City Clinical Hospital No. 67, Moscow Healthcare Department, 123423 Moscow, Russia
[8] Laboratory for Support of AI-Advised Clinical Decision Making, Sechenov First Moscow State Medical University, 119991 Moscow, Russia

e-mail: eshirshin@gmail.com

*Received May 12, 2023*
*Revised August 18, 2023*
*Accepted September 28, 2023*

One of the main sources of errors when conducting biochemical analysis of blood serum in a clinical diagnostic laboratory is the excessive concentration of hemoglobin (hemolysis) or lipids (lipemia) in the analyzed sample. Therefore, an important step to accurately determine the concentration of the target analyte is to first classify the sample into ”suitable” and ”unsuitable” classes for analysis. At the same time, to be used in practice, the method of preanalytical classification of samples must be both simple to implement and reliable, from the point of view of high sensitivity and specificity. In this work, we investigated the analytical ability of two approaches — an approach based on diffuse reflectance spectroscopy, characterizing the parameters of diffuse reflection of blood serum in the visible and near-IR range (500−1000 nm), and an approach based on computer vision — in classifying blood serum samples for normal suitable for analysis, and samples with hemolysis and lipemia. Diffuse reflectance spectroscopy has been found to demonstrate high sensitivity and specificity (more than 97%) in the classification of serum samples, but technically this method requires the application of a measuring probe to the sample. At the same time, computer vision methods have made it possible to determine the suitability of a sample for further analysis with lower classification accuracy values, but in more complex conditions, in particular, in the case of a sample moving along a conveyor line in a clinical diagnostic laboratory. The advantage of the studied methods, in addition to the high accuracy of preanalytical classification, is the simplicity of their technical implementation, as well as the ability to characterize samples without additional sampling of blood serum, which indicates their promise as methods for preanalytical analysis of blood serum samples.

**Keywords:** diffuse reflectance spectroscopy, computer vision, lipemia, hemolysis, blood serum, preanalytics.

DOI: 10.61011/EOS.2023.09.57353.5025-23

## Introduction

A biochemical blood analysis is one of the routine tests done in every clinical diagnostic laboratory. Most analyses of this kind are performed for blood serum after erythrocyte sedimentation by centrifugation: a laboratory assistant or an automated system take a serum from spun venous blood, mix it with an active agent, and read out the concentrations of analytes of interest, which is often done with the use of a photometric approach [1,2]. The validity of the obtained biochemical blood analysis result is crucial both for diagnosing and patient dismissal [3,4].

The majority of errors in laboratory testing arise at the preanalytical stage of the entire process [5]. The primary cause of errors (both false-positive and false-negative results) in biochemical analyses is the alteration of

normal optical properties of blood serum due to an excess concentration of hemoglobin (hemolysis), lipids (lipemia), or bilirubin and products of its degradation (icterus) in the sample. Hemolysis, lipemia, and icterus are commonly associated with procedural factors, such as incorrect blood sampling (e.g., non-fasting or careless sampling that leads to hemolysis) [6–9] and improper sample transport conditions [10], or individual problems of a patient (an unhealthy diet [11] or diseases leading to hemolysis and lipemia [12]).

The average percentage of samples unsuitable for analysis falls within the 3−5% range [13,14]. However, the number of samples with hemolysis, lipemia, and icterus may reach several tens of percent in certain wards or groups of patients; for example, the highest level of hemolysis (11.6%) and lipemia (1.6%) is observed empirically in blood samples taken from children under the age of three years [15], and the percentage of samples unsuitable for analysis in emergency departments varies from 6.8 to 19.8% [16]. The primary cause of unsuitability of samples in both outpatient and inpatient departments is hemolysis: hemolytic blood serum samples account for 40−70% of all unsuitable ones [17], while lipemic samples account for another 10−20% [15,18].

Hemolysis, lipemia, and icterus are easy to detect visually, since they alter the optical properties of blood serum. Specifically, hemoglobin features characteristic absorption lines in the blue-green spectral region (the Soret band at 410 nm and the Q 520−580 nm band) and gives a red color to blood serum. An increased concentration of lipids induces the formation of lipid droplets that make blood serum turbid and opaque, and icteric samples have characteristic yellow coloring attributable to the absorption of light by bilirubin (its absorption band is at 430−480 nm). The optical properties of blood serum may be evaluated (i.e., hemolysis, lipemia, and icterus may be identified) visually by a laboratory assistant, but this method is subjective and correlates only loosely with standardized automated preanalytical techniques; owing to this, such techniques are being used more and more often [19–21].

Automated methods for preanalytical classification of samples may, in turn, be split into groups of methods requiring the use of reagents (e.g., the cyanmethemoglobin method for measuring the hemoglobin concentration) and „unenhanced“ spectrophotometric methods that do not require additional sample preparation [22–25]. For example, an ADVIA2400 (Siemens) clinical chemistry system uses two wavelengths (571 and 596 nm) to estimate the hemoglobin content, two wavelengths (658 and 694 nm) to determine the degree of lipemia, and yet another two (478 and 505 nm) to estimate the bilirubin content [26].

However, not all laboratories with a low or moderate throughput are fitted with spectrophotometric preanalytical instruments. A separate solution for preliminary estimation of the degree of hemolysis and lipemia at the preanalytical stage may help in this case— specifically, a solution combining the possibility of visual inspection of a sample by a laboratory assistant with rapid identification of hemolysis, lipemia, and icterus performed simultaneously by a standardized approved instrument in a manner so that the assistant can manually lift a test tube to a scanning device without needing to open this tube and take a sample.

Since hemolysis and lipemia may be identified visually, computer vision techniques have been proposed as a means for non-contact evaluation in a number of studies. Both „classical algorithms“ of machine learning and computer vision [27] and deep learning techniques [28,29] have been used to identify hemolysis and lipemia. These methods have demonstrated high sensitivity and specificity in determination of the degrees of hemolysis and lipemia [28].

In the present study, we examine the possibility of preanalytical classification of samples into normal, hemolytic, and lipemic ones with the use of two optical techniques: diffuse reflectance spectroscopy, wherein a signal reflected from blood serum is analyzed at different optical and near-IR wavelengths, and computer vision methods (analysis of images of test tubes containing blood serum). In addition to requiring no sample preparation, these methods have an advantage in the ease of their engineering implementation and may be used both in semi-automatic (with a laboratory assistant analyzing each sample separately, which is suitable for small-scale laboratories) and automatic (e.g., in screening for hemolysis and lipemia on a conveyor line in a large clinical laboratory) modes. The obtained results indicate that both approaches have good prospects for application in detection of hemolysis and lipemia in samples.

## 1. Materials and methods

### 1.1. Studied samples

Venous blood samples in test tubes 4.9 ml in volume with a barrier gel (S-Monovette, Germany), which were subjected to centrifugation to separate the erythrocyte pellet from blood serum, were used to evaluate the analytical ability of diffuse reflectance spectroscopy and computer vision methods in preanalytical classification of serum samples into normal, hemolytic, and lipemic ones.

All measurements and manipulations with samples were performed in a certified accredited clinical diagnostic laboratory (ISO 15189:2012) of the Vorohobov City Clinical Hospital No. 67 (Moscow). Studies were approved by the local ethics committee of the hospital.

Blood serum samples with a hemoglobin concentration above 0.5 g/L („hemolysis“) and a concentration of lipids above 1.25 g/L („lipemia“) were selected for evaluation of the analytical performance of the examined techniques. These concentrations correspond to „soft“ thresholds and cover all semi-quantitative categories „+“ — „++++“ [15] that are often used for characterization of hemolysis and lipemia. The concentration of lipids and hemoglobin in „normal“ samples was below 1.25 and 0.5 g/L, respectively. The concentration of lipids and hemoglobin was verified

in a separate spectrophotometric study performed by measuring the optical density of serum with a Perkin-Elmer-25 spectrophotometer within the 300–800 nm range. No more than 1 ml of supernatant of spun blood were sampled into a quartz cuvette with an optical path of 2 mm for this purpose, and the optical density of the sample was determined. The hemoglobin concentration was determined by the oxyhemoglobin absorption line at 572 nm, and the concentration of lipids was determined at 660 nm by comparison with the optical density of lipofundin with a concentration of 1 g/L. Diffuse reflectance spectroscopy and computer vision tests were performed in the initial test tubes used for blood collection without additional sampling.

A total of 113 samples of spun blood were selected for examination by diffuse reflectance spectroscopy and with the use of computer vision algorithms: 33 samples of serum suitable for further biochemical analysis („normal"), 59 samples with varying degrees of lipemia, and 21 samples with varying degrees of hemolysis. Spun blood samples were stored in a refrigeration chamber at a temperature of 4°C. The measurement of spectra of diffusely reflected light and the acquisition of images to be processed with computer vision techniques in the manual mode were performed within 72 h of blood sampling. This ensured stability of whole blood samples and their validity for further hematological analysis [30].

A separate set of data on 150 samples of different classes was compiled for testing the computer vision algorithms in the „pipeline" mode (see below).

### 1.2. Measurement of diffuse reflectance spectra

A special dual-fiber diffuse reflectance detection setup was designed for experiments with blood serum samples in test tubes. A system of two convex lenses in this setup collimates light from a halogen lamp with a power of 15 W (with emission wavelengths falling within the 450–2000 nm range) on the end of an illuminating multimode silicon fiber (IPG Photonics, Russia) with a core diameter of 550 $\mu$m and a numerical aperture of 0.22. The other end was in contact with the measurement area of the examined sample. Another identical fiber with one of its ends positioned at distance $d = 0.2$ mm from the edge of the illuminating fiber was used to collect the diffusely reflected signal, which was then sent to the slit of a YSM-8101 spectrometer (Yixist, China) that detects radiation within the 200–1100 nm spectral range with a spectral resolution of $\sim$ 10 nm.

The diffuse reflection signal from blood serum samples was measured in the following way: the ends of illuminating and collecting fibers were pressed perpendicularly against the wall of a plastic test tube in the area corresponding to blood serum, and the reflected signal spectrum was then recorded.

In addition to intensity $I(\lambda)$ of the signal from a blood serum sample, response $I_{\text{ref}}(\lambda)$ of a reference Spectralon white sample (Labsphere, United States) with a reflectance of 99% in visible and infrared ranges and background signal $I_{bg}(\lambda)$, which included the detector noise, were measured in order to calculate diffuse reflectance spectrum $R(\lambda)$. This calculation was performed in accordance with the following formula:

$$R(\lambda) = \frac{I(\lambda) - I_{bg}(\lambda)}{I_{\text{ref}}(\lambda) - I_{bg}(\lambda)}, \quad (1)$$

and the diffuse reflectance was then converted into effective optical density $\text{OD}(\lambda)$ that was used to characterize the samples:

$$\text{OD}(\lambda) = -\ln R(\lambda). \quad (2)$$

Since the intensity of the detected signal below 500 nm was significantly lower than the intensity in the 500–1000 nm range, the diffuse reflectance signal within the spectral interval of 500–1000 nm was used for further analysis.

### 1.3. Photographic imaging of samples with a camera under standardized conditions

In order to perform an initial assessment of the analytical ability of computer vision methods applied to the classification of spun blood samples into normal, hemolytic, and lipemic ones, these samples were photographed under fixed lighting conditions.

A sample of spun venous blood to be photographed was positioned on a specialized support at a distance of $\sim$ 20 cm from a digital camera that produces RGB images with a resolution of1920 $\times$ 1080 pixels (Logitech C922 Pro, China). Imaging was performed in a light box insulated from ambient light; a LED white light source with a luminous flux of 4880 lm and a color temperature of 6220 K was used for illumination. The sample was positioned so that the area corresponding to blood serum (the color characteristics of which were to be evaluated) was within the field of view of the camera.

### 1.4. Photographic imaging of samples in the mode of operation with a conveyor line

A setup for preanalytical classification of samples with the use of computer vision algorithms in the high throughput mode (with blood serum samples moving along a conveyor line in a clinical diagnostic laboratory) was designed.

This setup included four digital cameras with an image resolution of 1920 $\times$ 1080 pixels (Logitech C922 Pro, China) spaced $\sim$ 30 cm apart and positioned at an angle of 90° to each other. These cameras imaged one and the same geometrical region from different sides simultaneously with a repetition rate of 15 Hz. A LED white light source (4880 lm and a color temperature of 6220 K) was installed above the camera mounts to provide uniform illumination of the area. The obtained images were sent to a control computer that stored and processed them.

The cameras were positioned in the following way: one of them was aligned with the direction of sample movement

along a conveyor line, another was directed oppositely, and the remaining two were set perpendicularly to the conveyor.

Data one 150 samples of each class („normal", „hemolysis," and „lipemia") were collected with the use of this setup. A total of 740 images of a test tube were selected for each class to train the model. To make it more robust, 76 photos containing no image of a test tube were added.

### 1.5. Labeling of images for classification and detection of test tubes

Manual labeling of selected images within Supervise.ly [31] was performed for the purpose of developing an algorithm fro detection and classification of test tubes. The process of manual data labeling within Supervise.ly consisted in highlighting the needed part of a test tube with a rectangular bounding box. In the present case, this rectangle extended from the end of the test tube holder to the top of blood serum. The sides of the box were snug against the boundaries of the object. The labeling result was a dataset containing the name of a photographic image, its size, and the class and coordinates of boundaries of the object in the image.

If the image area associated with blood serum was covered with a bar code glued onto a sample, only this bar code was highlighted, and the sample was classified as a one with a „hidden area."

### 1.6. Procedure of construction of classifiers and evaluation of their accuracy

A series of classifier models predicting whether a sample is normal, hemolytic, or lipemic were developed in order to evaluate the analytical ability of diffuse reflectance spectroscopy and computer vision methods in the mode with fixed conditions and the „pipeline" mode with a test tube moving along a conveyor line. Proprietary Python 3 scripts utilizing the NumPy, Pandas, Matplotlib, PyTorch, Ultralytics, and Scikit-Learn libraries [32–37] were used to construct and analyze these models. Let us describe the procedures of development and assessment of classification performance for each algorithm in more detail.

#### *1.6.1. Classifying and assessment of the quality of classification based on diffuse reflectance spectroscopy data.*

The model for classification of samples into normal, hemolytic, and lipemic ones was trained with 70% of the overall number of observations (optical density (OD) spectra), while the remaining 30% of measurements were used to evaluate the accuracy of the model (i.e., test it). A total of 30 random partitions of the dataset into training and test samples were generated to determine the errors of classification quality metrics.

The amplitudes of decomposition of OD spectra into the first three principal components, which were retrieved from the decompositions of spectra from the training sample, were used as sample classification criteria. A random forest algorithm with 100 trees without restrictions on the depth and the minimum number of objects in a tree leaf and the method of support vectors with a soft margin and a Gaussian kernel with a coefficient adjusted on the basis of accuracy for the validation sample were used to classify preprocessed data. The implementation of principal components analysis (PCA), the random forest algorithm, and support vectors from the scikit-learn library [37] was used.

#### *1.6.2. Classifying and assessment of the quality of classification with the use of computer vision in fixed sample positions.*

Images obtained under standardized lighting conditions in fixed test tube positions were processed by highlighting their areas corresponding to blood serum. The average values of intensity in R (red), G (green), and B (blue) color channels, which correspond to the spectral ranges of 600−680, 520−600, and 430−500 nm, respectively, were calculated for these regions. These values then served as predictors for classification of test tubes.

As in Section 1.6.1, random forest and support vector methods were used for classification. It was determined empirically that the model based on the method of support vectors with a soft margin and a Gaussian (RBF) kernel with a coefficient of 0.33 has the best configuration. The cross-validation technique was used for assessment. The training set contained 70% of data, and the remaining 30% were included into the validation set. The training procedure was repeated 100 times with premixing and selection of new data for training and validation sets.

#### *1.6.3. Assessment of the quality of detection by the computer vision model for classification of samples on a conveyor line.*

A model searching for an object in an image (highlighting the image area corresponding to the object) and classifying this object (i.e., solving the problem of detection) was needed for operation in the high-throughput mode with a sample moving along a conveyor line. The YOLO v5 model [36] was used for the purpose.

The search performance was evaluated using the mean average precision (mAP) metric with the degree of overlap between ground-truth and predicted object areas being above 50% (intersection over union ¿ 50%). This metric characterizes the ratio between the number of correctly identified objects and the overall number of detector actuations in the case when the degree of overlap between ground-truth and predicted areas exceeds 50%.

„Hidden area" (the blood serum area is covered by a bar code glued onto the wall of a test tube) and „background" (no object in the image) classes were added into the model alongside with the „normal" „hemolysis," and „lipemia" classes mentioned above.

Eighty percent of the sample were used for training, and the remaining 20% served for validation. The YOLO v5 architecture was initialized with weights pretrained on the

COCO dataset [38] for better convergence and fine-tuned on the training dataset with the use of stochastic gradient descent. Prior to application of the model, images were converted to a size of $980 \times 980$ pixels; the model was fine-tuned with 50 epochs of data with a batch size of 2.

#### *1.6.4. Calculation of classification quality metrics.*

Classifiers were characterized by ROC curves of the dependence of sensitivity and specificity of a classifier on the threshold of labeling an object as the one belonging to a positive class. The sensitivity and specificity of joint identification of hemolysis and lipemia were calculated in the following way: hemolytic and lipemic samples were assumed to belong to a positive class, normal samples were assigned to a negative class, and the sensitivity and specificity were then evaluated in accordance with the standard procedure for binary classification. Metrics were calculated with the use of scikit-learn [37] built-in implementations.

## 2. Results

### 2.1. Assessment of diffuse reflectance spectroscopy in preanalytical classification of blood serum

The possibility of application of diffuse reflectance spectroscopy in identification of hemolysis and lipemia in blood serum samples in the reflection geometry (using just the diffusely reflected signal and without the need to open a test tube) was examined first. A detection setup with two optical fibers with a core diameter of 550 $\mu$m spaced 0.2 mm apart was used for the purpose. These fibers were brought close to a visually observable area of a blood serum sample, and light from a broadband radiation source (halogen lamp emitting within the 450–2000 nm range) was transmitted along one of them. The reflected signal was collected by the other fiber that sent it to a spectrometer operating within the 200–1000 nm range (see the description of this setup in Section 1.2). Figure 1, *a* presents the schematic diagram of measurement of diffuse reflectance spectra from a certain area of a blood serum sample in the dual-fiber setup.

Samples with hemolysis and lipemia were selected for evaluation of sensitivity of this classification approach: the minimum concentration of hemoglobin and lipids in these samples was 0.7 and 1.3 g/L, respectively. According to the guidelines for preanalytical laboratory diagnostics, this corresponds to a low degree of hemolysis and lipemia in blood serum samples.

Diffuse reflectance spectra were measured for all samples, and the effective OD was calculated based on the obtained data as a negative logarithm of reflectance (see formula (2)). Figure 1, *b* shows the class-averaged mean effective OD spectra for normal, lipemic, and hemolytic samples. It can be seen that normal samples had low OD values within the 600–900 nm interval, while the absolute OD values of lipemic samples in the near IR range were much higher. This behavior of optical density is attributable to the fact that samples with marked lipemia have an enhanced scattering factor due to the presence of lipid droplets in the solution (Fig. 1, *b*). At the same time, samples with hemolysis did not feature high OD values in the IR range, but lines typical of oxyhemoglobin absorption were observed within the 500–600 nm interval in the visible range. The low intensity of the radiation source at wavelengths below 500 nm and a strong hemoglobin absorption in certain samples interfered with the detection of the diffuse reflectance signal in this spectral region. Such features of effective OD spectra allow one to group samples easily into normal, lipemic, and hemolytic classes. It was found that the samples of different classes are easy to distinguish by applying PCA to OD spectra in planes corresponding to the first three components (Fig. 1, *c*).

A simple classification algorithm was constructed for quantitative evaluation of the accuracy of classification of samples into normal, hemolytic, and lipemic ones. Effective OD spectra within the 500–1000 nm range served as the input data for this algorithm, which then transformed them via PCA in such a way that amplitudes of the first three principal components of OD spectra were used as feature vectors of samples instead of the initial spectral data. Following dimensionality reduction, a random forest model or a model based on the method of support vectors were used to classify samples in accordance with their spectral features (amplitudes of three principal components). Since the random forest model was found to be the most accurate in an evaluation performed using a validation sample, the results for this algorithm are reported below.

The classification accuracy was evaluated in a standard way: 70% of objects from the entire sample were used to train the model (both to select principal components and train the classification algorithm), and the remaining 30% were used to evaluate the accuracy of the classifier. A total of 30 random partitions of the dataset into training and test samples were generated to determine the errors of classification quality metrics.

Sensitivity and specificity metrics (the proportions of correctly identified truly positive samples and correctly identified truly negative samples) are often used in classification problems. However, the values of sensitivity and specificity may depend on the threshold at which the probability of belonging to a certain class is converted into a predicted class label. In view of this, the performance of a classifier is evaluated with the use of a ROC curve (dependence of the sensitivity on specificity at each value of the mentioned classification threshold). The area under a ROC curve (AUC-ROC) may act as an „integral estimate“ of a classifier. This area may actually be interpreted as an average value of classification sensitivity at all levels of specificity. Figure 1, *d* presents the ROC curves for the problem of identification of lipemic, hemolytic, and normal samples. It can be

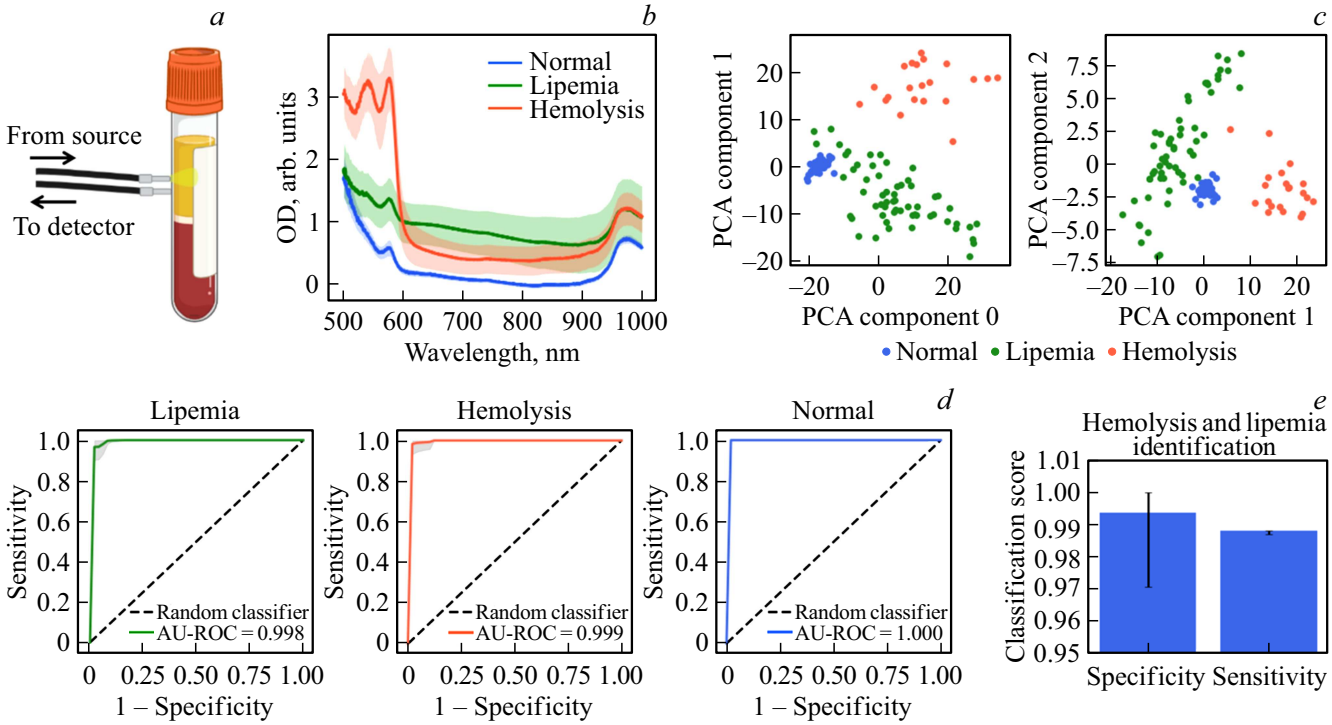


**Figure 1.** (*a*) Schematic diagram of the method for diffuse reflectance measurements for a sample of spun blood serum in the dual-fiber setup. (*b*) Mean and intragroup mean-square deviation for effective OD spectra of blood serum with marked lipemia, samples with hemolysis, and samples suitable for biochemical analysis („normal“). (*c*) Result of application of PCA to OD spectra dimensionality reduction (amplitudes of the first three principal components are indicated) in testing of blood serum samples. (*d*) ROC curves for classification of samples into lipemic, hemolytic, and normal ones plotted with the use of cross-validation. (*e*) Cross-validation data on the sensitivity and specificity in identification of lipemia and hemolysis in samples with the use of the constructed classifier.

seen that the classification accuracy is close to unity for all classes.

The sensitivity and specificity in detection of hemolysis and lipemia among all samples have also been evaluated (Fig. 1, *e*). It was found that the discussed classifier has a sensitivity of 99% at a specificity above 99.4% (according to cross-validation data). The obtained results suggest that this method has good prospects for application in detection of hemolysis and lipemia in blood serum. It is also conceivable that this approach will be used for quantitative and semi-quantitative characterization of samples: determination of the range of concentrations of hemoglobin and lipids in blood serum.

Since three diffuse reflection parameters (amplitudes of three principal components) were found to be sufficient for classification, we decided to examine the possibility of classification of samples based on reflection data with the use of computer vision algorithms. Experiments in „ideal conditions“ with images of fixed test tubes under constant illumination were performed first in order to evaluate the classification accuracy, and the accuracy of the computer vision model in „real-world conditions“ (in classification of samples moving along a conveyor line in a clinical diagnostic laboratory) was investigated after that.

### 2.2. Assessment of accuracy of preanalytical classification of samples with the use of computer vision: model conditions

It was proposed to examine the color contrast of images obtained under fixed lighting conditions in order to evaluate the accuracy of preanalytical classification of test tubes with samples of spun blood serum studied by diffuse reflectance spectroscopy (Section 2.1). In addition to fixed lighting conditions, the test tube was always positioned in the same area of obtained images, simplifying significantly the task of highlighting the area corresponding to blood serum. These experiments were aimed at determining the „upper limit“ of accuracy of preanalytical classification of samples based on color characteristics of blood serum with the use of computer vision under such conditions when all external factors (first and foremost, illumination and test tube position) remain unchanged. An example labeled image of a test tube with spun blood is presented in Fig. 2, *a*.

Figure 2, *b* presents the examples of highlighted areas of blood serum of various classes („normal“, „lipemia„, and „hemolysis“). Following the discussed segmentation procedure, the average values of intensity in the red (R), green (G), and blue (B) image channels were calculated for areas of blood serum. Figure 2, *c* shows the scatter diagrams for average intensity values in different color channels for

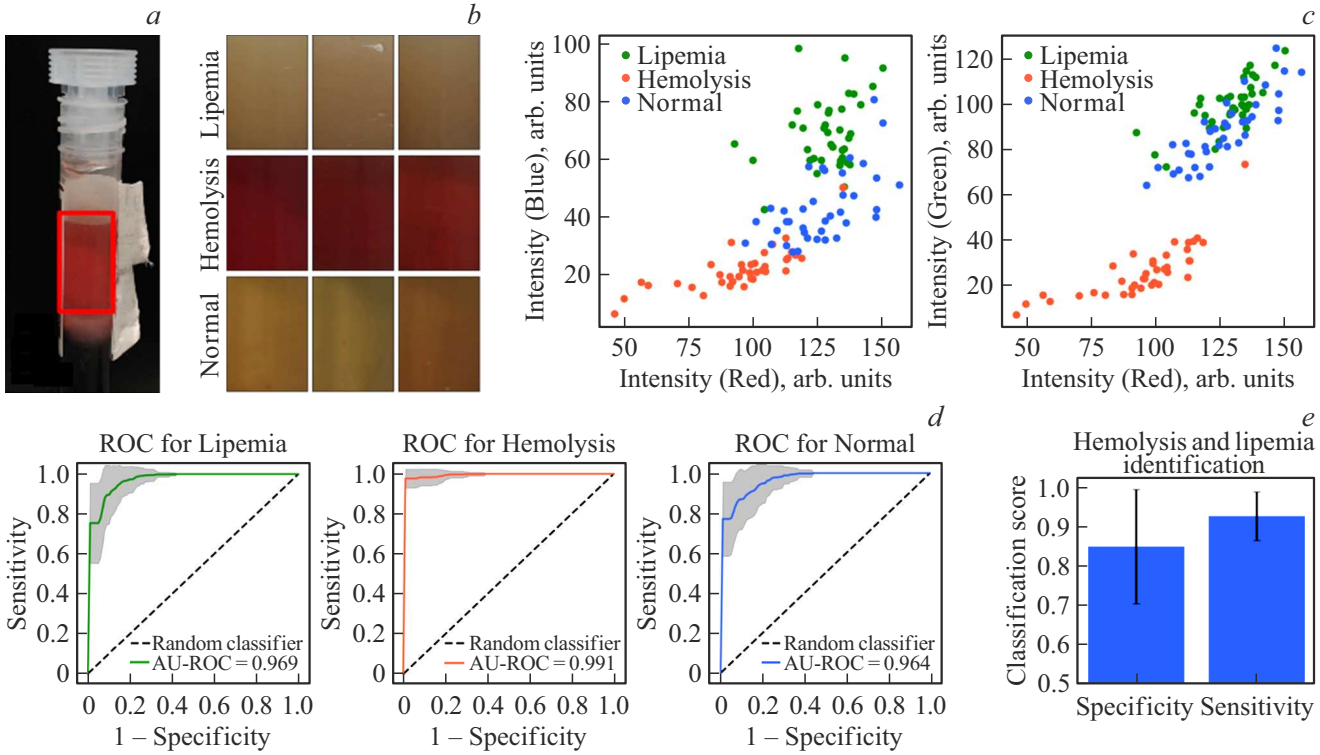


**Figure 2.** (*a*) Example image of a test tube with blood serum with marked hemolysis. The red rectangle highlights the area corresponding to blood serum. (*b*) Examples of highlighted blood serum areas in samples belonging to different classes (normal, lipemia, and hemolysis). (*c*) Scatter diagrams for average signal intensity values in the R, G, and B image channels for segmented areas of blood serum. (*d*) ROC curves for the classification of samples into lipemic, hemolytic, and normal ones based on the RGB values of blood serum with the use of a support vector classifier. (*e*) Sensitivity and specificity of identification of hemolysis or lipemia. The classification threshold was chosen in such a way as to maximize the classification sensitivity.

these areas. It can be seen that dots corresponding to different classes (normal, lipemia, and hemolysis) occupy non-overlapping domains. Therefore, a high accuracy of preanalytical classification based on the average intensity values in the R, G, and B image channels is to be expected.

Several classification algorithms based on the random forest and support vectors methods, which accepted average intensity values in the R, G, and B channels as features and predicted whether a sample belongs to „normal“, „lipemia„, or „hemolysis“ classes, were then developed. It was found empirically that a support vector machine model with a soft margin and a Gaussian kernel produced the most accurate results for the validation sample. The sensitivity and specificity of classification were evaluated for this classifier at all threshold levels (see the ROC curves in Fig. 2, *d*). The areas under ROC curves for all three classes fell within the range of 0.96–0.99 (a unit area corresponds to an „ideal“ classifier). The sensitivity and specificity of identification of hemolysis and lipemia with the use of color characteristics of blood serum were also determined: the maximum sensitivity was $90.9 \pm 7.5\%$ at a specificity of identification of hemolysis and lipemia of $84 \pm 15\%$. It is evident that the sensitivity and specificity values fluctuate significantly between different partitions of the dataset. This may be attributed to the presence of „outliers“ in the training sample in certain partitions. That said, the average values of sensitivity and specificity of preanalytical classification with the use of computer vision are also high, although diffuse reflectance spectroscopy has an advantage in this regard.

### 2.3. Assessment of accuracy of preanalytical classification of samples with the use of computer vision: testing in the pipeline mode

Imaging and image analysis have an advantage in imposing potentially less strict requirements regarding the positioning of an object relative to a camera than other methods (specifically, diffuse reflectance spectroscopy) that produce accurate results only if the experimenter mounts the probe in direct proximity to analyzed blood serum. Owing to this, computer vision methods may be applied in certain advanced cases (e.g., when the position of a test tube in an image is not fixed or when a test tube moves relative to the camera). This formulation of the problem may be relevant to large laboratories where samples move along a conveyor line: having installed a preanalytical analyzer on a conveyor line, one may screen the samples additionally for hemolysis and lipemia and discard unsuitable ones.

A setup designed for a conveyor line was constructed in order to apply computer vision algorithms in this formulation of the problem. The setup included four digital cameras spaced $\sim 30$ cm apart and positioned at an angle of 90° to each other. These cameras imaged one and the same geometrical region from different sides simultaneously with a frame rate of 20 fr/s. Four cameras were needed because of the fact that a test tube with a sample may have an adhesive label with an identification bar code covering the „useful" part of blood serum within the field of view of a single camera. Cameras were positioned in immediate proximity to an additional LED white light source that established uniform illumination conditions for samples moving along the conveyor line. The obtained images were sent to a control computer that processed them. A photographic image of the discussed module mounted on the conveyor line is shown in Fig. 3, *a*. The cameras of this module were positioned in the following way: two oppositely directed cameras were mounted parallel to the trajectory of test tube movement along the conveyor line, while the remaining two were set perpendicularly to the conveyor. Example images made by a camera aligned with the movement of a sample and a camera directed perpendicularly to the conveyor are shown in Fig. 3, *b*. This prototype unit was used to compile a dataset for samples of spun blood.

Images from four cameras were recorded for blood serum samples, and several frames for each test tube were selected to be used in the development of computer vision models for preanalytical classification of test tubes. A total of 3 776 images taken by different cameras were collected for training of the developed algorithm. The training sample was balanced so as to include 185 images of each class (normal, hemolysis, and lipemia) taken from four different angles, 185 images of test tubes with the blood serum region covered by an identification bar code (additional class „hidden area"), and 76 additional images without test tubes, which are needed to enhance the accuracy of operation of the computer vision model. Images of test tubes in the database compiled for training of the model were labeled: the image area occupied by a test tube was highlighted (rectangles in Fig. 3, *b*), and the class of a test tube was indicated. Examples of segmented areas belonging to different classes are presented in Fig. 3, *c*.

Let us describe the procedure of development of a computer vision model in more detail. In the current configuration of the detection setup, a test tube with a sample may be positioned anywhere within the image. Therefore, additional models „searching for" a test tube and classifying it were needed to detect objects and perform their preanalytical classification. This problem may be solved in two stages: the position of an object (test tube) in an image is identified first, and analysis is performed after that (e.g., analysis of the color of blood serum; see Section 2.2. above). However, a single-stage detection approach, wherein the same model is used to search for an object in an image and classify it, is more robust and was proven useful in such applications. This is the reason why we chose this approach implemented with the use of a YOLOv5 [36] convolutional neural network that identifies the indicated four classes of samples („normal", „hemolysis", „lipemia," and „hidden area") and indicates the position of objects in an image with rectangular bounding boxes. Training and test samples did not include images corresponding to one and the same test tube. Twenty percent of images from the conveyor line were used for validation, and the other 80% were used for training.

The quality of detection of an object in an image was characterized with a common mean average precision (mAP) metric [39]. It characterizes the ratio between the number of correctly identified objects and the overall number of detector actuations in the correct area. A value of $\mathrm{mAP}_{@\mathrm{IoU}=0.5} = 0.984$ was obtained in the case when the degree of overlap between ground-truth and predicted areas was 50% ($\mathrm{IoU} = 0.5$). Thus, it may be concluded that the model identified fairly accurately the position of a test tube in photographic images.

The average values of sensitivity and accuracy of the classifier were monitored in order to examine the variation of accuracy of classification with model parameters. It was found that the average classification accuracy is 0.964 at a class determination sensitivity of 0.972. Figure 3, *d* shows the confusion matrix of the classifier: the distribution of true class labels over classes predicted by the developed model. The prediction thresholds were set so as to maximize the sensitivity of detection of hemolytic and lipemic samples.

It follows from the confusion matrix that the model adjusted this way identifies hemolytic samples with a unit accuracy and identifies lipemic samples with an accuracy of 0.94, but occasionally (in 23% of cases) produces false-positive identifications of hemolysis in normal samples. If the blood serum area is not visible in an image („hidden area" class), a sample is never classified as a normal, hemolytic, or lipemic one. When a test tube was not highlighted during labeling („background" class), the model distributed objects evenly among all classes. However, according to the $\mathrm{mAP} = 0.984$ metric, such images were produced rarely (their occurrence rate was less than 2%) when test tubes were imaged by a camera from a considerable distance.

The overall sensitivity and specificity of identification of hemolysis and lipemia among samples with visible blood serum (i.e., among „normal", „hemolysis", and „lipemia" classes only) were also evaluated: the sensitivity of identification of hemolysis and lipemia was 97% at a specificity of 74% (Fig. 3, *e*). Low specificity levels are attributable primarily to those cases when the algorithm erroneously classified samples with normal blood serum as hemolytic ones. The algorithm may be improved further by compiling a larger observation dataset, optimizing the parameters of the algorithm, and including images taken from different angles under various lighting conditions into the training sample.

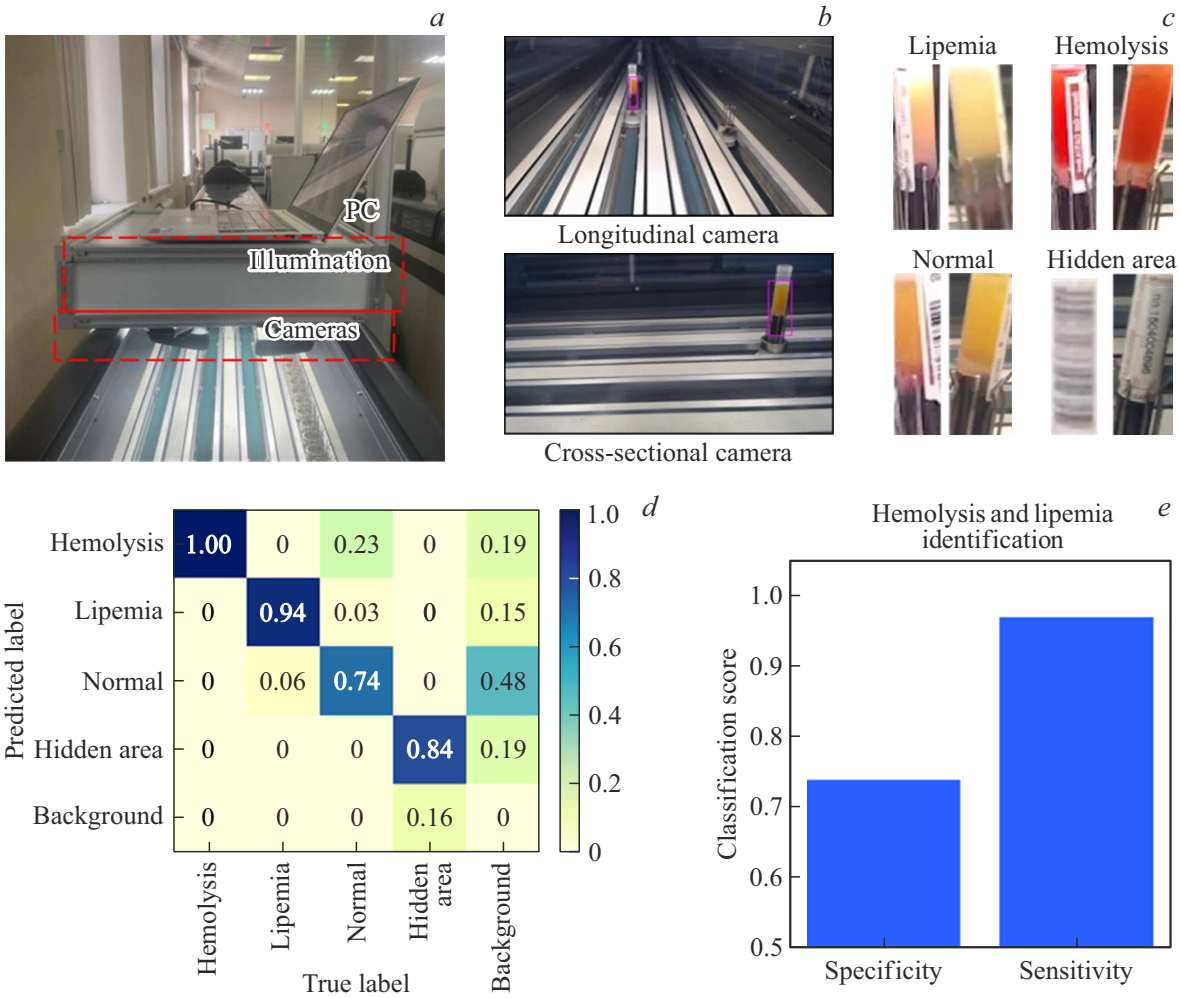


**Figure 3.** (*a*) Photographic image of the prototype setup that takes images of samples on the conveyor line in a clinical diagnostic laboratory. (*b*) Example images from cameras aligned with the movement of a test tube (top) and directed perpendicularly to the conveyor (bottom). (*c*) Examples of segmented areas of images of different sample classes („normal“, „hemolysis“, „lipemia,“ and „hidden area“). (*d*) Confusion matrix for the problem of preanalytical classification with the use of the developed computer vision model. Numbers denote the proportion of objects with the corresponding true class value. (*e*) Sensitivity and specificity in identification of hemolysis and lipemia among samples.

## 3. Discussion

The performance of two methods for preanalytical classification of venous blood serum (diffuse reflectance spectroscopy and computer vision) applied to the problem of classification of samples into normal, hemolytic, and lipemic ones was evaluated. These methods have an advantage in being contactless: measurements may be performed directly through the glass wall of a test tube if blood serum is clearly visible.

Diffuse reflectance spectroscopy turned out to be the most accurate of the examined methods. Data for 113 samples of blood serum with varying degrees of hemolysis and lipemia and normal samples were processed, and a high sensitivity (above 99%) at an average specificity of 99.4% was determined for the classifier that uses the spectra of effective optical density (negative logarithm of diffuse reflectance) within the 500–1000 nm interval as features. Note that the method of principal components, which converts the initial OD spectra into amplitudes of the first three principal components, was applied to the input spectral data for dimensionality reduction. In other words, the constructed classifier used only three „spectral“ features within the indicated wavelength range to classify samples. Therefore, instead of recording broadband diffuse reflectance spectra, one may isolate a small number of wavelengths (3–4 spectral bands) for measurement of diffuse reflectance and classify samples with the same efficiency, eventually making it unnecessary to use a spectrometer. The addition of measurements performed at shorter wavelengths (400–500 nm) may help make the method more robust in regard to identification of samples with an elevated hemoglobin concentration and to detection of an elevated bilirubin concentration based on its absorption band with a

maximum at $\sim 460$ nm. It is also worth mentioning that samples were classified into normal, hemolytic, and lipemic ones at low thresholds of concentration of lipids (1.25 g/L) and hemoglobin (0.5 g/L) in blood serum. Setting several thresholds for these concentrations, one may perform semi-quantitative characterization of the degree of lipemia and hemolysis of a sample.

However, the proposed fiber method for measurement of diffuse reflectance spectra requires that a probe be positioned in direct proximity to the sample area containing blood serum. This makes the method ill-suited for automatization. Therefore, in our view, this technique has niche applications in small-scale laboratories where automated preanalytical studies are not performed and samples are examined manually. Diffuse reflectance spectroscopy then helps eliminate subjectivity and the factor of human error in evaluation of the properties of blood serum.

The accuracy of preanalytical classification of blood serum samples with the use of computer vision algorithms was evaluated next. The results of classification of samples into normal, hemolytic, and lipemic ones with the test tube position and the level of illumination being fixed were reported in Section 2.2, and Section 2.3 is focused on a more complex model that simultaneously searches for a sample and classifies it in the case when this sample is moving along a conveyor line in a medical laboratory.

It was found that average values of intensity in the R, G, and B channels for images areas corresponding to blood serum provide an opportunity to perform accurate classification with a hemolysis and lipemia detection sensitivity of $91 \pm 7.5\%$ at a specificity of $84 \pm 15\%$ in the first case (with the sample position fixed and under constant illumination). False-positive identifications of hemolysis or lipemia in normal samples were distributed uniformly over the „hemolysis“ and „lipemia“ classes. These false positives may be excluded by adjusting the probability threshold at which the classifier labels a sample as the one belonging to a certain class.

A more complex model based on a YOLOv5 neural network was developed for detection and classification of samples in a more advanced case (when they move along a conveyor line). It was found that this model has a low proportion of false positives in searching for samples within an image: $\mathrm{mAP50} = 0.984$, which corresponds to 1.6% of false positives). If the test tube area corresponding to blood serum was visible, the maximum hemolysis and lipemia identification sensitivity for samples detected in an image was 97% at a specificity of 74%. The majority of false positives were those occurring when a normal sample was classified as a hemolytic one. This is presumably attributable to the fact that when samples move along a conveyor line, the level of illumination of the field within which test tubes are detected is less uniform, affecting the reproduction of color of blood serum.

This hypothesis is supported, e.g., by the observation that the model is sensitive to the level of illumination of the fields of view of cameras and the field of view of a camera: when these parameters changed, the accuracy of detection of objects and their classification decreased. Owing to this, the final data were collected in the same measurement configuration. The model may be made more robust in regard to illumination levels by fine-tuning it with images corresponding to various illumination conditions and different camera positions; however, this requires a proportionally greater number of images to be included into the training dataset. In spite of its mentioned shortcomings, the computer vision approach is potentially applicable in preanalytical classification of blood serum. More specifically, the applicability of this method in identification of samples with hemolysis and lipemia on a conveyor line in a medical laboratory was demonstrated.

## Conclusion

Preanalytical screening of blood serum samples for suitability for further biochemical analysis is a crucial step in establishing a correct diagnosis. In the present study, the analytical ability of diffuse reflectance spectroscopy and computer vision algorithms in preanalytical screening of samples for hemolysis and lipemia, which are major causes of errors in biochemical analyses, were examined. It was found that diffuse reflectance spectroscopy in a dual-fiber setup with measurements performed without additional sampling directly through the wall of a test tube provides high (above 99%) levels of sensitivity and specificity in identification of samples with hemolysis (with a hemoglobin concentration above 0.7 g/L) and lipemia (with a concentration of lipids above 1.3 g/L). Computer vision methods may also be used to identify hemolysis and lipemia. Specifically, the analysis of color characteristics (intensity levels in standard color channels) of blood serum imaged with a camera under fixed lighting conditions offers a selectivity in excess of $97 \pm 2.5\%$ at a specificity of identification of hemolysis and lipemia of $89 \pm 8\%$. „Advanced“ computer vision techniques (convolutional neural networks) are applicable in real-time analysis of images of test tubes moving relative to a detector. However, the sensitivity of a model depends strongly in this case on lighting conditions, the positioning of a test tube with respect to a detector, and the quality of obtained images.

### Acknowledgments

This study was carried out as part of the „Priority 2030“ academic leadership program of the Sechenov First Moscow State Medical University of the Ministry of Health of the Russian Federation (Sechenov University) and as part of scientific activities of the Interdisciplinary Research and Education School „Photonic and Quantum Technologies. Digital Medicine“ of the Moscow State University.

### Funding

This study was supported by a grant from the Government of Moscow for applied medical research project No. 2212-19/2.

### Conflict of interest

The authors declare that they have no conflict of interest.

## References

[1] J.Z. Ji, Q.H. Meng. Clinica Chimica Acta, **412** (17–18), 1550–1553 (2011).
[2] M.B. Smith, Y.W. Chan, A. Dolci, M.D. Kellogg, C.R. McCudden, M. McLean. *Wayne, PA, USA: Clinical and Laboratory Standards Institute* (2012).
[3] E.P. Kakorina, A.V. Polikarpov, N.A. Golubev. Laboratory Service, **7** (4), 32–39 (2018).
[4] P.L. Epner, J.E. Gans, M.L. Graber. BMJ Quality & Safety, **22** (Suppl 2), ii6–ii10 (2013).
[5] P. Bonini, M. Plebani, F. Ceriotti, F. Rubboli. Clinical Chemistry, **48** (5), 691–698 (2002).
[6] G. Lippi, G.L. Salvagno, G. Lima-Oliveira, G. Brocco, E. Danese, G.C. Guidi. Clinica Chimica Acta, **440**, 164–168 (2015).
[7] A.M. Simundic, K. Bölenius, J. Cadamuro, S. Church, M.P. Cornes, E.C. van Dongen-Lases, P. Eker, T. Erdeljanovic, K. Grankvist, J.T. Guimaraes, R. Hoke. Clinical Chemistry and Laboratory Medicine (CCLM), **56** (12), 2015–2038 (2018).
[8] G. Lima-Oliveira, G. Lippi, G.L. Salvagno, M. Montagnana, G. Picheth, G.C. Guidi. Biochemia Medica, **22** (2), 180–186 (2012).
[9] F. Sanchis-Gomar, G. Lippi. Biochemia Medica, **24** (1), 68–79 (2014).
[10] G. Lippi, G.L. Salvagno, E. Danese, G. Lima-Oliveira, G. Brocco, G.C. Guidi. Clinica Chimica Acta, **436**, 183–187 (2014).
[11] A.-M. Simundic, M. Cornes, K. Grankvist, G. Lippi, M. Nybo. Clinica Chimica Acta, **432**, 33–37 (2014).
[12] W. Barcellini. Transfusion Medicine and Hemotherapy, **42** (5), 287–293 (2015).
[13] A. Abdollahi, H. Saffar, H. Saffar. North Am. J. of Medical Sciences, **6** (5), 224 (2014).
[14] G. Lippi, M. Plebani, A.-M. Simundic. Biochemia Medica, **20** (2), 126–130 (2010).
[15] G. Tian, Y. Wu, X. Jin, Z. Zeng, X. Gu, T. Li, X. Chen, G. Li, J. Liu. J. Plos One, **17** (1), e0262748 (2022).
[16] N.J. Heyer, J.H. Derzon, L. Winges, C. Shaw, D. Mass, S.R. Snyder, P. Epner, J.H. Nichols, J.A. Gayken, D. Ernst, E.B. Liebow. Clinical Biochemistry, **45** (13–14), 1012–1032 (2012).
[17] G. Lippi, N. Blanckaert, P. Bonini, S. Green, S. Kitchen, V. Palicka, A.J. Vassault, M. Plebani. Clinical Chemistry and Laboratory Medicine, **46** (6), 764–772 (2008).
[18] R. Chawla, B. Goswami, D. Tayal, V. Mallika. Laboratory Medicine, **41** (2), 89–92 (2010).
[19] M.R. Glick, K.W. Ryder, S.J. Glick, J.R. Woods. Clinical Chemistry, **35** (5), 837–839 (1989).
[20] G. Lippi, J. Cadamuro, A. von Meyer, A.M. Simundic. Clinical Chemistry and Laboratory Medicine (CCLM), **56** (5), 718–727 (2018).
[21] A.M. Simundic, N. Nikolac, V. Ivankovic, D. Ferenec-Ruzic, B. Magdic, M. Kvaternik, E. Topic. Clinical Chemistry and Laboratory Medicine, **47** (11), 1361–1365 (2009).
[22] D.A. Noe, V. Weedn, W.R. Bell. Clinical Chemistry, **30** (5), 627–630 (1984).
[23] C. Bürki, M. Volleberg, D. Brunner, M. Schmugge, M. Hersberger. Clinical Biochemistry, **100**, 67–70 (2022).
[24] S. Storti, E. Battipaglia, M.S. Parri, A. Ripoli, S. Lombardi, G. Andreani. J. Laboratory Medicine, **43** (2), 67–76 (2019).
[25] C.-J.L. Farrell, A.C. Carter. Ann. Clin. Biochem., **53** (5), 527–538 (2016).
[26] Z. Du, J. Liu, H. Zhang, B. Bao, R. Zhao, Y. Jin. J. Clin. Lab. Anal., **33** (4), e22856 (2019).
[27] Z. Wang, Z. Zhao. In: *MATEC Web Conf. 173*, EDP Sciences (2018).
[28] H. Wang, H. Huang, X. Wu. Chemom. Intell. Lab. Syst., **231**, 104688 (2022).
[29] C. Yang, D. Li, D. Sun, S. Zhang, P. Zhang, Y. Xiong, M. Zhao, T. Qi, B. Situ, L. Zheng. Clin. Chim. Acta, **531**, 254–260 (2022).
[30] M. Ashenden, A. Clarke, K. Sharpe, G. d'Onofrio, J. Plowman, C.J. Gore. Int. J. Lab. Hematol., **35** (2), 183–192 (2013).
[31] Supervise.ly [Electronic source].
URL: https://supervisely.com/
[32] C.R. Harris, K.J. Millman, S.J. Van Der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern. Nature, **585**, 357–362 (2020). DOI: 10.1038/s41586-020-2649-2
[33] W. McKinney. In: Proceedings of the 9th Python in Science Conference, **445** (1), 51–56 (2010).
[34] J.D. Hunter. Computing in Science & Engineering, **9** (03), 90–95 (2007).
[35] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison. Advances in Neural Information Processing Systems, **32** (2019).
[36] Ultralytics [Electronic source].
DOI: 10.5281/zenodo.7347926
[37] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas. J. machine Learning Research, **12**, 2825–2830 (2011).
[38] T.Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick. In: *Computer Vision–ECCV 2014: 13th European Conferenc Proceedings*, *Part V 13* (2014).
[39] R. Padilla, S.L. Netto, E.A. Da Silva. In: *International conference on systems, signals and image processing* (*IWSSIP*), 237–242 (2020).

*Translated by D.Safin*