На основі експериментально встановлених даних щодо параметрів мікроструктури, елементного та фракційного складу порошків титанових сплавів виділено чотири класи їх відповідності як вихідної сировини для адитивних технологій – матеріал з відмінними властивостями, оптимальними властивостями, в матеріалі можливі дефекти і бракований матеріал. Встановлено основні характеристики матеріалу, які визначають його приналежність до певного класу. Для реалізації процедур машинного навчання побудовано навчальну та тестову вибірки на основі 20 ознак, які характеризують кожен з чотирьох класів порошків титанових сплавів. Описано розроблений метод ідентифікації класу матеріалу, який базується на сумісному використанні поліному Колмогорова-Габора другого степеня та алгоритмі Випадкового лісу. На основі точності роботи в режимах навчання та застосування проведено експериментальне порівняння результатів роботи розробленого методу з результатами роботи існуючих методів: Випадкового лісу, Логістичної регресії та Машини опорних векторів. Наведено візуалізацію результатів роботи усіх досліджених методів.

Розроблений метод керованого навчання дозволяє будувати моделі опрацювання великої кількості ознак кожного вхідного вектора. При цьому алгоритмом Випадкового лісу забезпечуються задовільні генералізуючі властивості при збереженні переваг додаткового підвищення точності на основі полінома Колмогорова-Габора.

Експериментально встановлено основні переваги розробленого методу, зокрема щодо додаткового підвищення точності розв'язання задачі класифікації. Розроблений метод дозволяє підвищити точність моделювання на 34,38; 33,34 та 3,13 % порівняно з методами: Машина опорних векторів, Логістична регресія та Випадковий ліс відповідно.

Отримані результати дозволяють значно скоротити фінансові та часові витрати під час виготовлення виробів методами адитивних технологій. Застосування інструментарію штучного інтелекту дозволяє зменшити трудомісткість та енерговитратність експериментів з визначення оптимальних характеристик порошкових матеріалів

Ключові слова: порошки титанових сплавів, мікроструктура, морфологія, гранулометричний склад, адитивні технології, методи штучного інтелекту

DEVELOPMENT OF MACHINE LEARNING METHOD OF TITANIUM ALLOY PROPERTIES IDENTIFICATION IN ADDITIVE TECHNOLOGIES

R. Tkachenko
Doctor of Technical Sciences, Professor,
Head of Department*
E-mail: roman.tkachenko@gmail.com

Z. Duriagina
Doctor of Technical Sciences, Professor**
Katolicki Uniwersytet Lubelski Jana Pawła II
Al. Racławickie, 14, Lublin, Poland, 20-950
E-mail: zduriagina@ukr.net

I. Lemishka
Postgraduate student**
E-mail: mzihor@ukr.net

I. Izonin
PhD, Assistant*
E-mail: ivanizonin@gmail.com

A. Trostianchyn
PhD, Assistant**
E-mail: trostianchyn@gmail.com
*Department of Publishing Information
Technologies***
**Department of Applied Materials Science
and Materials Engineering***
***Lviv Polytechnic National University
S. Bandery str., 12, Lviv, Ukraine, 79013

1. Introduction

The present stage of the material science development is characterized by the accumulation of a large amount of experimental data on the relationship between the structure and properties of materials of various functional purposes [1–3], which operate under different exploitation conditions [4, 5]. This, in turn, causes the designing of new, unique in properties materials. The processes of development, technological processing, approbation, and implementation are long-lasting, costly and complex [6, 7].

Under such conditions, the sustainable development of material science involves several possible approaches:

– the analysis of existing information accumulated over the years about certain research objects, in particular, by machine learning methods (ML) [8] and extrapolation of the obtained results;

– the investigation the "samples of witnesses" of new materials by traditional methods with the transferring of generalized results to "reference" samples.

The combination of traditional experimental investigation of the materials with using models, methods, and tools

of artificial intelligence [9, 10] increases the efficiency of the procedure for developing or designing new materials. Traditional approaches allow obtaining all necessary information about the material properties [7], and the usage of powerful [11], modern ML algorithms [12–13] makes this process easier, shorter and cheaper. This is due to solving prediction and regression tasks, classification or clustering on a small experimental data sample and extrapolation of the results to the new material.

That is why the development of new and modification of existing approaches for designing and constructing new materials based on such approach is the actual task.

2. Literature review and problem statement

In many cases, traditional approaches to establishing a relationship between the structure and properties of materials are commonly used and justify themselves [1, 2]. However, when it comes to the significant non-linearity of such relationships, it is difficult to establish the necessary parameters. Particularly acute is the problem of multi-parameter dependencies processing. It is these tasks that cause the search for new, more effective ways to obtain the required information.

Artificial intelligence tools usage to solve material science tasks; is not sufficiently widespread, but interest to this area is steadily increasing. In the review papers [8, 14], some methods of artificial intelligence in the supervised and unsupervised modes, including machine learning algorithms are described. These methods have obvious advantages for data pre-processing using Principal Component Analysis [14]. However, they are not tools for solving regression or prediction tasks [14].

In [15], the Artificial Neural Networks (ANN) were used for solving the polystyrenes temperature prediction task. It should be noted that such an approach for prediction and classification tasks has a number of disadvantages, the main of which is that most of the ANN paradigms do not provide a repeatability of the solution. In [16], the solution of the task for predicting the reaction outcomes for the crystallization of template vanadium selenites on the basis of the Support Vector Machine (SVM) usage is described. The solution of the classification problem in the field of material science, based on the use of the same tools, is given in [8]. The main disadvantage of the proposed methods is the need for the correct selection of optimal parameters, in particular, the kernel, to ensure an effective result. In [17], a method for solving the regression task to determine the high-performance metalorganic frameworks for CO_2 capture was developed. SVM with the radial-basis function core was chosen as a tool for solving this task. Based on the high speed of the SVM and the high accuracy provided by the radial-basis function, the method shows satisfactory results. However, the method is sensitive to the data standardization and noises.

In [8], the solution of the prediction task of the material's friction coefficients based on Decision trees is described. The advantages of these algorithms, both in prediction and regression tasks, and in classification tasks are the high speed and accuracy of their work. However, the methods of this class require the correct selection of optimal parameters for their work. Otherwise, they occupy a huge amount of memory, which limits their practical use. In addition, this tool shows poor results on noisy data.

The literature review shows the perspectives of machine learning algorithms application for solving various tasks of material science (forecasting, detection of anomalies, classification, recognition, regression). However, the main problem is the right choice of machine learning algorithm. On the one hand, it should ensure high accuracy of work, and on the other hand, it should be quick and easy, not require large computing resources and computer memory.

Known methods do not always provide a combination of the above characteristics, which imposes a number of limitations on their practical application. In this regard, an important task is to improve the existing and develop new methods of artificial intelligence in the field of material science.

3. The aim and objectives of the study

The main aim of this work is to develop the method for the identification of the titanium alloys conformity based on the parameters of microstructure and properties of their powder fractions, by solving the classification task using means of machine learning.

To achieve the aim, it is necessary to accomplish the following objectives:

– to conduct experimental studies on the determination of the microstructure and properties of titanium alloy powders of different fractions and micro geometry of the surface, to form training and test samples;

– to apply the Kolmogorov-Gabor polynomial and the Random Forest algorithm for solving the classification task. To investigate such composition of the method in terms of the accuracy during the identification of the object of research while minimizing time resources required to implement the training procedures;

– to establish experimentally the machine learning method parameters, which would provide the optimal result in terms of time and accuracy of its work;

– to compare the accuracy of the proposed method with existing ones and to develop recommendations for its application.

4. Inputs data and research methods

4. 1. Investigation of the material properties

Spherical powders of titanium alloys are obtained by the technology of centrifugal atomization of an electrode [18], and non-spherical – by the hydrogenation-dehydrogenation technology [19].

The formation of the inputs database for the implementation of the machine learning process was carried out on the basis of studies of morphology, elemental and granulometric composition of titanium alloy powders of various systems [20, 21] previously conducted by the authors and based on the usage of literary sources [22, 23].

The morphology of the particle surface structure was studied using a scanning electron microscope EVO 40XVP.

The distribution of powders by fractions was performed by the sieve method in accordance with DSTU ISO 565: 2007. To evaluate the micro geometry of the surface structure, as well as the extended fractional analysis of the powders, a software product ImageJ was used to analyze the microstructures [24]. The degree of inhomogeneity (polydis-

persity) of the powder, which depends on the average size of the dominant particles in a certain fraction and the standard deviation of the particle size of the powder from the average size, was determined by the construction of the Gaussian curve based on the histogram of the particle distribution in a certain fraction [25].

Artificial intelligence tools can be used to reduce the duration, as well the cost of investigation of the properties of spherical and nonspherical titanium alloy powders [10]. The conducted literature review has shown the feasibility of using the algorithms of machine learning to solve this task, in particular on the basis of the Random Forest algorithm and the Kolmogorov-Gabor polynomial.

4. 2. Random Forest algorithm

The machine learning algorithms based on Decision trees since their creation and to this day have been given great attention. Such methods are used for solving applied problems in various areas.

The algorithm for constructing one binary Decision tree works according to the scheme of the Greedy algorithm. During each iteration, the hypersurface partition of class space is constructed for the input set of training sample vectors, which minimizes the average measure of the heterogeneity between two obtained subsets. This procedure is performed recursively for each received subset until the criteria for stopping are met [26].

Obviously, for constructing a model based on only one Decision tree, you can get a solution that is sensitive to noise. Therefore, it is expedient to use an ensemble of several trees, which is typical for the Random Forest algorithm. This leads to the number of advantages, including:

– possibility to get a stable and effective solution to the task by combining responses from each tree;

– a multi-tree ensemble avoids problems associated with the method overfitting, or at least minimizes them [27];

– an independent training procedure for each individual tree from the ensemble at its sample's part provides the possibility to apply methods of this class in distributed computing systems, in particular using methodology [28].

In order to solve the classification task, the Random Forest algorithm uses a large number of trees, each of which is taught in a separate subset of the set of vectors of the entire training sample. The response of each tree from the ensemble is taken into consideration in the following way. Each of them is noted for the affiliation of the type of conformity, then they are averaged, and based on their largest number, the winner is determined.

However, averaging is not always the most effective option for assembling Random forest. There are others [29], such as Simple Voting, Weighted Voting, Mixture of Experts, and so on. However, their use is limited to tasks that require the implementation of a fast training procedure.

4. 3. Kolmogorov-Gabor polynomial

The Kolmogorov-Gabor polynomial is often used as an effective tool for approximating multi-parameter dependencies. This tool allows modeling with very high accuracy in the generalization mode [30].

In this case, the polynomial degree plays an important role. In case of increasing the polynomial degree, the approximation possibilities for dependencies with essential nonlinearity are improved, however, the data generalization properties are worsened. The experimental investigations carried out with available data allowed choosing the second-degree Kolmogorov-Gabor polynomial as a model of optimal complexity. This polynomial can be written as follows:

$$Y(x_1,...,x_n) = ai_0 + \sum_{i_1=1}^{n} ai_1 \times xi_1 + \sum_{i_1=1}^{n} \sum_{i_2=i_1}^{n} ai_1 i_2 \times xi_1 \times xi_2, \quad (1)$$

where n – the number of variables for each data vector.

The procedure for finding the coefficients of this polynomial is a non-trivial task. Existing methods are rather complex or very lengthy. In addition, they do not always provide sufficient accuracy of the result, in particular, when constructing a system of equations with initial conditionality, or in the case of a significant correlation of the inputs, or when the task is almost degenerate.

5. Results of the initial database formation and computer modeling

5. 1. Initial database formation

Fig. 1 shows the morphology of the investigated spherical and non-spherical powders of systems Ti-6Al-4V and Ti-Al-V-Zr.

It should be noted that the main characteristics by which the investigated powders will be classified into certain classes (excellent material properties, optimal material properties, possible defects in the material and defective material) (Fig. 2), are the average diameter of the powder particles (Fig. 3) and polydispersity (Fig. 4), which are partially taken from [31, 32].
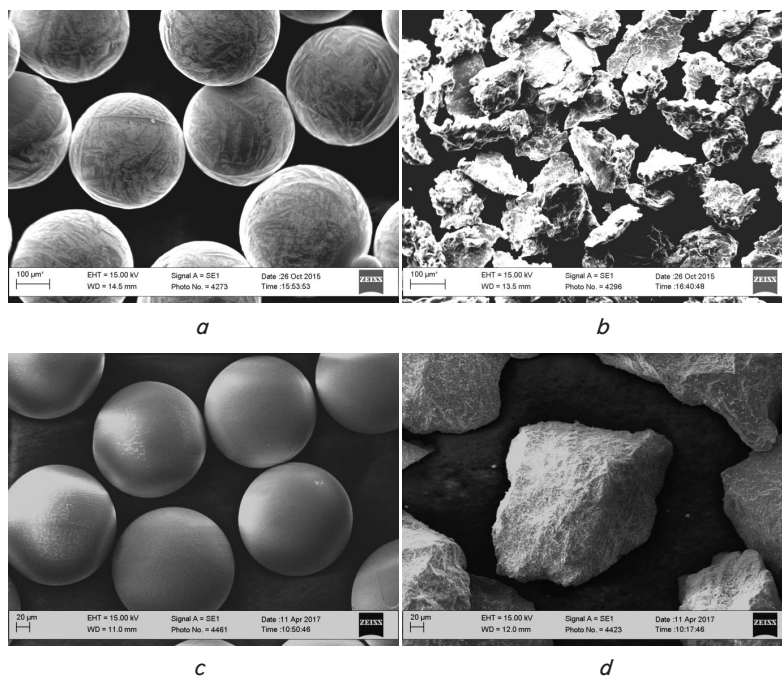

Fig. 1. Morphology of titanium alloy powders:
a, b — system of Ti-6Al-4V [18]; *c, d* — system of Ti-Al-V-Zr [19]
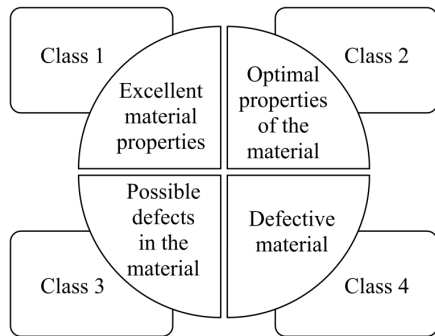
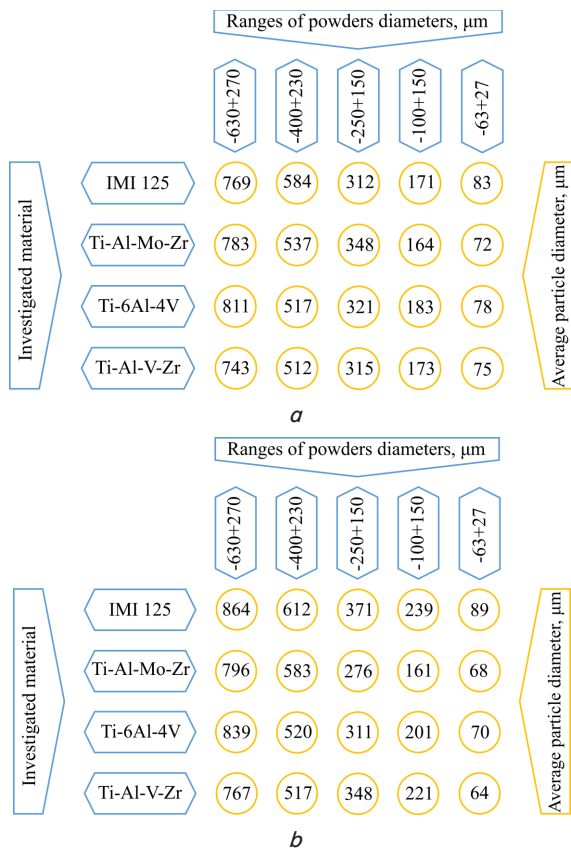Fig. 2. Classes of conformity of powders of the investigated titanium alloys

The algorithmic implementation of the method involves the creation of a database, which in our case consists of 480 vectors, each of which contains 20 input characters (Fig. 5) [10]. These attributes determine belonging to one of the four classes of material mentioned above (Fig. 4).
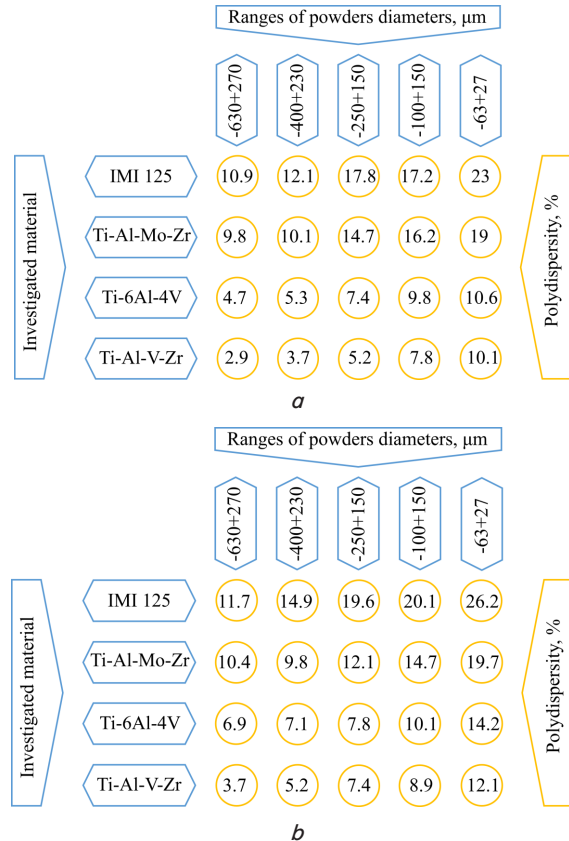

Fig. 3. The average diameter of powder particles of the investigated titanium alloys:
a — spherical; *b* — non-spherical

Fig. 4. Polydispersity of powder particles of the investigated titanium alloys: *a* — spherical; *b* — non-spherical

In the case of processing images with a large number of small details, the geometrical image super-resolution methods to accelerate the evaluation of all the required parameters of the investigation material can be used.

5. 2. Sample formation for the implementation of machine learning procedures

The data sample to solve the task posed in the work based on the experimental data was formed.
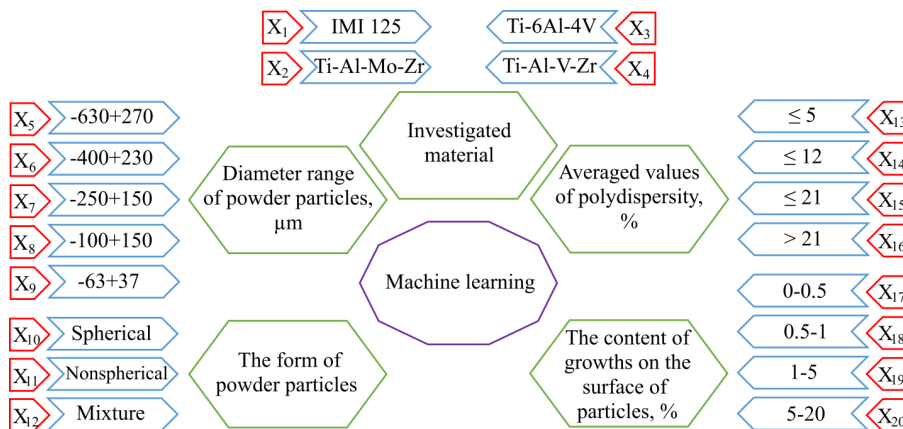

Fig. 5. Schematic representation of the source database for modelling

In Fig. 6, in 2D space, 4 studied classes of material (Fig. 2) based on 20 input characteristics (Fig. 5) are visualized. For this purpose, the method of machine learning FreeViz using the Orange software (version 3.8.0.) [33] was used. Different colors and shapes mark the characteristics of four different classes. In order to better visualize the input data, in Fig. 6, *b*, the results of visualization after the optimization procedure by the above method are presented. It should be noted that the optimization procedure was conducted solely for the purpose of a clearer visual representation of different classes.
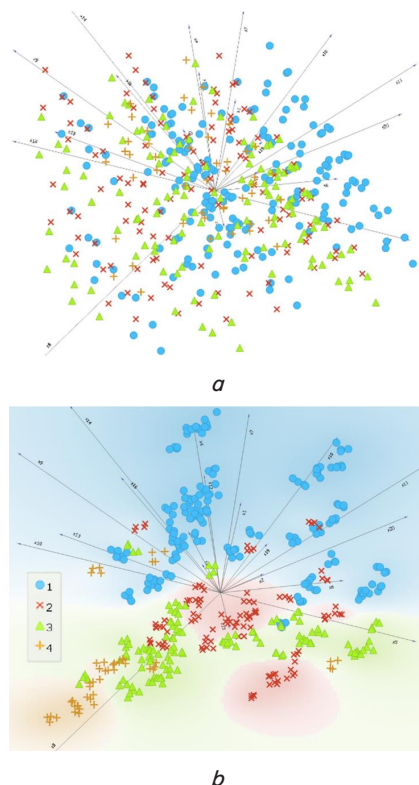

a

b

Fig. 6. 2D visualization of 4 material classes with 20 input features: *a* — before optimization; *b* — after optimization

As can be seen in Fig. 6, *b*, the experimental data are grouped in the colored regions, each of which represents a separate powder material class. The blue area separates the first class, that is, the case when the material is characterized by excellent properties. The red area denotes the second class of material, which has the optimal characteristics. The green area combines the data obtained for the material in which the defects are detected. It is not recommended for creating important parts. The last, fourth, or orange area groups data with the characteristics of the defective material that is not generally recommended for use.

480 vectors of experimental data, each of which is characterized by 20 characteristics, were randomly divided into a training and test sample. The ratio of such division is 80 and 20 % respectively. Fig. 7 shows histograms of the quantitative representation of data vectors of each of the four classes that were used during modeling for both types of samples.

As can be seen from Fig. 7, the most representative for both samples is the first class (the case where the material is characterized by excellent properties) and the least presented is the fourth study class of material that is not recommended for use.
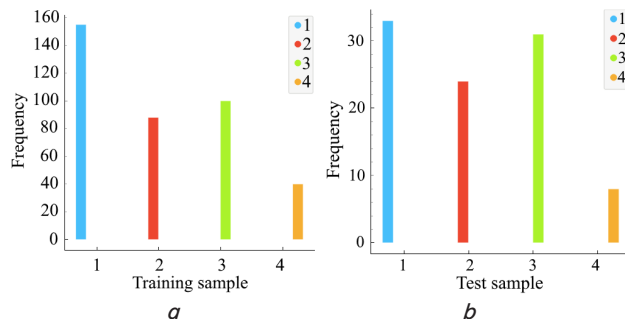

a   *b*

Fig. 7. The ratio of data vectors for each of four classes: *a* — training sample; *b* — test sample

5. 3. Composition of the proposed classification method

In the paper, the combined use of the Kolmogorov-Gabor polynomial and the Random Forest algorithm to increase the accuracy of the solution of the material class identification task based on the proposed characteristics is proposed. Input characteristics from Fig. 5 ($x_1...x_{20}$) of each vector are represented as the polynomial members, according to (1). The Random Forest algorithm is used to find the coefficients for Kolmogorov-Gabor polynomials. The expediency of using the developed method for solving this task is due to the following assertions:

– it provides an opportunity to effectively process a large number of characteristics of each input vector (the number of features during their representation in the form of Kolmogorov-Gabor polynomial members significantly increases);

– it allows one to efficiently work with small data samples while providing a sufficient level of generalization, and with large ones, with a minimization of the probability of overfitting.

Based on the obtained coefficients, the Kolmogorov-Gabor polynomial is used, which, having high approximation properties provides a high-precision result.

5. 4. Computer modeling for material class identification

The simulation of the material class identification method was carried out on the software developed by the authors. A number of libraries of the programming language Python [35] have been used for this purpose. The main parameters of the developed method are as follows:

– the degree of the polynomial was 2;

– the number of trees in the Random Forest algorithm was 9;

– the minimum number of objects at which splitting is performed was 2;

– the maximum depth of trees;

– the splitting criterion was classic, Gini index.

The accuracy of the method was estimated by the number of correctly classified samples to the dimension of the test sample in percentage terms.

It is found that the accuracy of the developed method at these parameters is 96.88 %.

6. Discussion of the developed method results

The combined use of the Kolmogorov-Gabor polynomial and the Random Forest algorithm ensures:

– on the one hand, sufficient generalization properties for constructing effective training models;

– on the other hand, saving the benefits of further increase *of* the accuracy of the result.

Both advantages are important in view of the expected high costs of creating a new material with unsatisfactory properties in case of a false identification. This, in turn, will negatively affect the performance of the design of aerospace equipment, the parts of which are planned to be made of the investigated materials by the 3D printing method.

Since increasing the number of trees in the Random Forest algorithm increases both the accuracy and the working time of the developed method, it is necessary to investigate the minimum number of trees that would provide the best result. Fig. 8 shows an experimental comparison of the accuracy of training and testing of the developed method when changing the number of trees in the machine learning algorithm under other equal conditions.
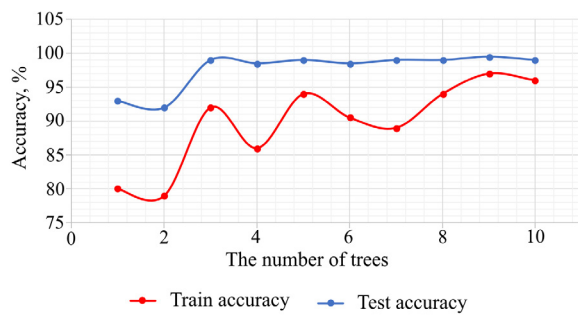

Fig. 8. Influence of the number of trees of the developed method the accuracy of its work

As can be seen from Fig. 8, the highest accuracy of the method is obtained with the least possible number of trees – 9. Apart from the fact that the specified method parameter provides the highest accuracy, it also provides the smallest difference between the training and testing accuracy, which is also can be seen from Fig. 8.

Comparison of the developed method results was carried out with the results of existing ones:
– Random Forest;
– Logistic regression;
– Support Vector Machine.

Table 1 presents experimental results of the modeling of all methods based on the accuracy both in the training and test modes.

Table 1

Comparison results of the developed method with the existing ones

Method	Train accuracy, %	Test accuracy, %
Developed method	100	96.88
Random Forest	100	93.75
Logistic Regression	78.85	63.54
SVM	76.5	62.5

As can be seen from Table 1, the best results, i.e. 96.88 % of accuracy, are obtained using the developed method. The worst results for solving the classification task are obtained by two well-known methods – the Support Vector Machine and the Logistic Regression.

Let us consider the modeling results in more detail. Fig. 9 provides a visual assessment of the work in the form of mosaic displays and scatter plots using Orange software (version 3.8.0.) [33]. The scatter plots provide visual information on how many class members are misidentified, and the mosaic displays clearly demonstrate which of these classes are assigned to these samples. Fig. 9, *a* shows the initial conditions for both diagrams, i.e. the ideal case where all samples are correctly identified.

The width of the mosaic display columns indicates the number of representatives of one or another class. The numerical ratio of this indicator can be found in Fig. 7. The scatter plot depicts 4 material classes, diagonally, starting from one. The *x*-axis is responsible for the true values of classes and the *y*-axis is responsible for the obtained values by one of the described methods. For a better visual perception of data, parameter *Jittering*=1 % of the Orange software environment is used to estimate the number of each class elements. In fact, all members of each separate class are overlaid and are at one point.

Information from both charts is important for this task. As already mentioned above, the accuracy of classification plays an important role, in particular in terms of the cost for developing the material, especially in the case of improper operation of the method. Incorrectly classified samples can cause large losses in the case when the part created based on such material will quickly fail and may cause the entire device to fail. Nevertheless, it is equally important to identify a class for which an incorrectly classified sample will be assigned. For example, from Fig. 9, *d* (results of the SVM work) it can be seen that three members of class 3 (the material with a defect) are identified as class 1, (the material with excellent characteristics), and one member of class 4 is classified as a class of material with optimal properties (class 2). This result is critical because based on such identification, material that is not recommended to be used for developing important parts can be used as having excellent characteristics. The same is the case with the application of the Logistic Regression method. Two samples from the class 3, i.e. the material with a defect, is classified as belonging to the class 1, and one sample from class 1 (specific material that is not recommended for use) is identified as a class with optimal characteristics.

Even in spite of the low classification results using the SVM and Logistic Regression, the above results are inadmissible because they can affect the adoption of an incorrect decision that will have negative consequences. It is not recommended to use such methods to solve this task.

Analyzing the developed method (Fig. 9, *c*), and the most similar to it – Random Forest (Fig. 9, *b*), it is possible to note the following. These methods identified respectively one and two samples of class 4 as class 3, three and one sample of class 3 (respectively for the Random Forest and the developed method) as class 2. Both methods classify one sample of the material class with optimal properties as the material class, which is characterized by excellent properties.

Such results are satisfactory from the standpoint of possible minor losses due to the incorrect identification, since situations where material with the defect is identified as the material with excellent properties, etc., do not have the place in the application of these methods. As for the accuracy of the results, the developed combination of the use of the Kolmogorov-Gabor polynomial and the Random Forest algorithm shows significantly better results in comparison to the classic Random Forest algorithm. Therefore, this method can be used for solving practical tasks of material science, which are critically sensitive to the accuracy of the result.
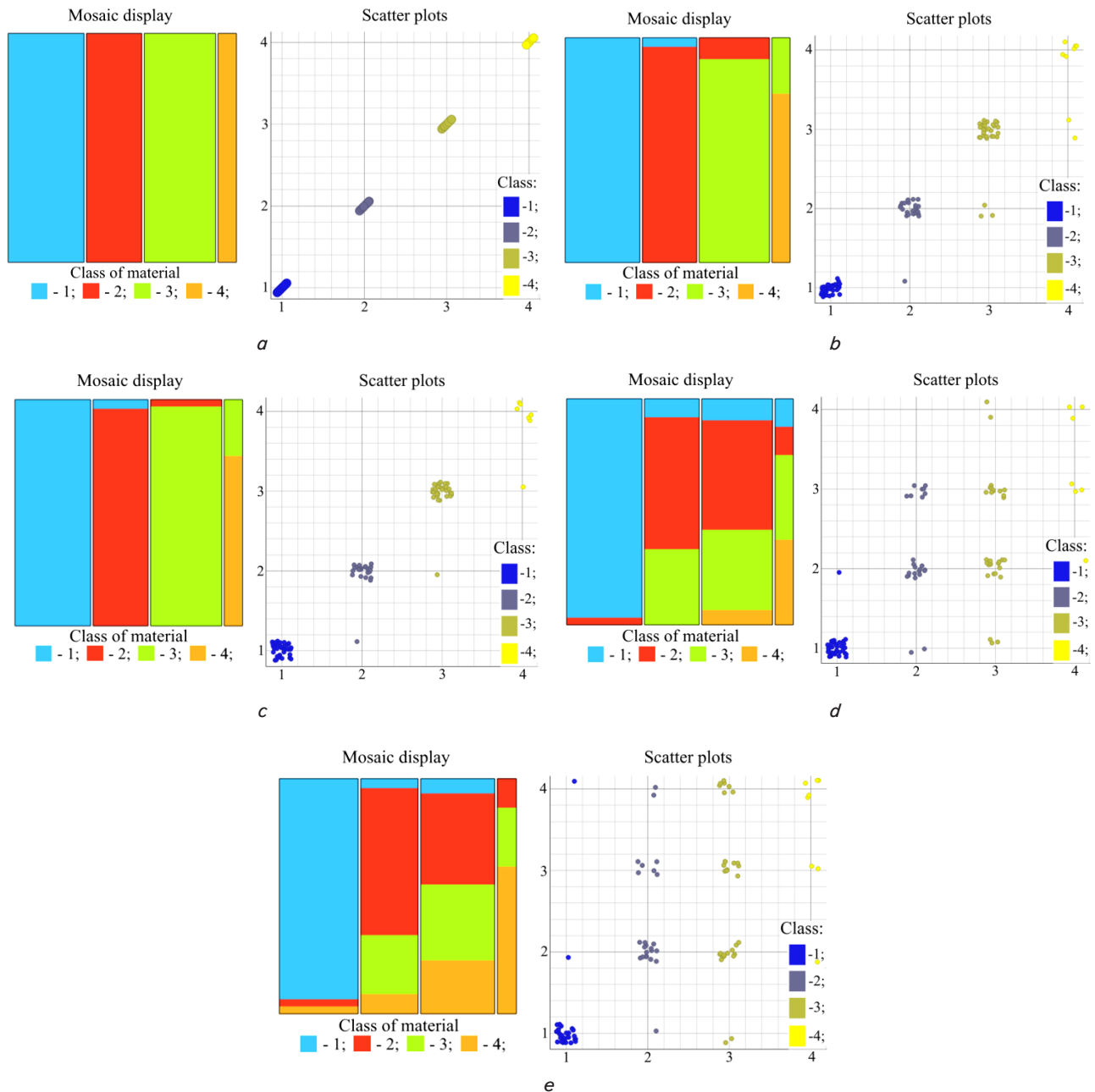
Fig. 9. Comparison of the methods based on the work results visualization:
a — Initial data; *b* — Random Forest; *c* — developed method; *d* — SVM; *e* — Logistic Regression

Further research can be conducted in the direction of applying new splitting criteria for the developed method. In order to improve the accuracy of the classification and clustering, regression, and prediction tasks for the solution of various problems in material science, it is also planned to use neural-like structures of the Successive Geometric Transformations Model [11].

7. Conclusions

1. Based on experimental data on the titanium alloy powders properties, 20 characteristics of their belonging to a certain class of raw material for the additive technologies have been identified. This allowed constructing training and test samples for the implementation of machine training procedures for the purpose of powder materials classification according to the parameters of microstructure, elemental and fractional composition.

2. A new classification method based on the combined use of the random forest algorithm and the Kolmogorov-Gabor polynomial has been developed. It was found that such a combination provided high accuracy of the result of solving the classification task – 96.88 %.

3. The expediency of using the developed method is confirmed by an experimental comparison of the results with existing methods. It is found that the developed method allows increasing the modeling accuracy by 34.38, 33.34 and 3.13 % compared with the methods: Support Vector Machine, Logistic Regression, and Random Forest respectively.

References

1. Duriagina Z. A., Kovbasyuk T. M., Bespalov S. A. The Analysis of Competitive Methods of Improvement of Operational Properties of Functional Layers of Flat Heating Elements // Uspehi Fiziki Metallov. 2016. Vol. 17, Issue 1. P. 29–51. doi: 10.15407/ufm.17.01.029

2. Pidkova V. Structure and properties of Mg, Al, Ti oxide and nitride layers formed by ion-plasma sputtering // Functional materials. 2015. Vol. 22, Issue 1. P. 34–39. doi: 10.15407/fm22.01.034

3. Magnetometric analysis of surface layers of 12X18H10T steel after ion-beam nitriding / Duryagina Z. A., Bespalov S. A., Borysyuk A. K. Pidkova V. Ya. // Metallofizika i Noveishie Tekhnologii. 2011. Vol. 33, Issue 5. P. 615–622.

4. Influence of the mode of thermal treatment and load ratio on the cyclic crack-growth resistance of wheel steels / Ostash O. P., Andreiko I. M., Kulyk V. V., Babachenko O. I., Vira V. V. // Materials Science. 2009. Vol. 45, Issue 2. P. 211–219. doi: 10.1007/s11003-009-9177-4

5. Influence of braking on the microstructure and mechanical behavior of railroad wheel steels / Ostash O. P., Andreiko I. M., Kulyk V. V., Vavrukh V. I. // Materials Science. 2013. Vol. 48, Issue 5. P. 569–574. doi: 10.1007/s11003-013-9539-9

6. On the concept of selection of steels for high-strength railroad wheels / Ostash O. P., Anofriev V. H., Andreiko I. M., Muradyan L. A., Kulyk V. V. // Materials Science. 2013. Vol. 48, Issue 6. P. 697–703. doi: 10.1007/s11003-013-9557-7

7. Fatigue crack growth resistance of welded joints simulating the weld-repaired railway wheels metal / Ostash O. P., Kulyk V. V., Poznyakov V. D., Haivorons'kyi O. A., Markashova L. I., Vira V. V. et. al. // Archives of Materials Science and Engineering. 2017. Vol. 2, Issue 86. P. 49–52. doi: 10.5604/01.3001.0010.4885

8. Mueller T., Kusne A. G., Ramprasad R. Machine Learning in Materials Science // Reviews in Computational Chemistry. 2016. P. 186–273. doi: 10.1002/9781119148739.ch4

9. Basic Components of Neuronetworks with Parallel Vertical Group Data Real-Time Processing / Tsmots I., Teslyuk V., Teslyuk T., Ihnatyev I. // Advances in Intelligent Systems and Computing. 2017. P. 558–576. doi: 10.1007/978-3-319-70581-1_39

10. Determination of the best microstructure and titanium alloy powders properties using neural network / Duriagina Z. A., Tkachenko R. O., Trostianchyn A. M., Lemishka I. A., Kovalchuk A. M., Kulyk V. V., Kovbasyuk T. M. // Journal of Achievements in Materials and Manufacturing Engineering. 2018. Vol. 87, Issue 1. P. 25–31. doi: 10.5604/01.3001.0012.0736

11. Imbalance Data Classification via Neural-Like Structures of Geometric Transformations Model: Local and Global Approaches / Tkachenko R., Doroshenko A., Izonin I., Tsymbal Y., Havrysh B. // Advances in Intelligent Systems and Computing. 2018. P. 112–122. doi: 10.1007/978-3-319-91008-6_12

12. Narushynska O., Teslyuk V., Vovchuk B.-D. Search model of customer's optimal route in the store based on algorithm of machine learning A // 2017 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT). 2017. doi: 10.1109/stc-csit.2017.8098787

13. Shakhovska N. B., Noha R. Y. Methods and Tools for Text Analysis of Publications to Study the Functioning of Scientific Schools // Journal of Automation and Information Sciences. 2015. Vol. 47, Issue 12. P. 29–43. doi: 10.1615/jautomatinfscien.v47.i12.30

14. Rajan K. Materials informatics // Materials Today. 2005. Vol. 8, Issue 10. P. 38–45. doi: 10.1016/s1369-7021(05)71123-8

15. Prediction of glass transition temperatures for polystyrenes from cyclic dimer structures using artificial neural networks / Xu J., Zhu L., Fang D., Liu L., Xu W., Li Z. // Fibers and Polymers. 2012. Vol. 13, Issue 3. P. 352–357. doi: 10.1007/s12221-012-0352-0

16. Machine-learning-assisted materials discovery using failed experiments / Raccuglia P., Elbert K. C., Adler P. D. F., Falk C., Wenny M. B., Mollo A. et. al. // Nature. 2016. Vol. 533, Issue 7601. P. 73–76. doi: 10.1038/nature17439

17. Rapid and Accurate Machine Learning Recognition of High Performing Metal Organic Frameworks for CO_2 Capture / Fernandez M., Boyd P. G., Daff T. D., Aghaji M. Z., Woo T. K. // The Journal of Physical Chemistry Letters. 2014. Vol. 5, Issue 17. P. 3056–3060. doi: 10.1021/jz501331m

18. Zlenko M., Popovich A., Mutylina I. Additivnye tekhnologii v mashinostroenii. Sankt-Peterburg, 2013. 221 p.

19. Haznaferov M. V., Ovchinnikov A. V., Yanko T. B. Tekhnologiya polucheniya «low-cost» poroshkov legirovannogo titana dlya additivnyh processov // Titan. 2015. Issue 2. P. 31–36.

20. Hranulometrychni kharakterystyky poroshku tytanovoho splavu VT20 otrymanoho metodom vidtsentrovoho plazmovoho rozpylennia elektrodu / Duriahina Z. A., Trostianchyn A. M., Lemishka I. A., Skrebtsov A. A., Ovchynnykov O. V. // Metaloznavstvo ta obrobka metaliv. 2017. Issue 1. P. 45–51.

21. The influence of chemical-thermal treatment on granulometric characteristics of titanium sponge powder / Duriagina Z., Trostyanchyn A., Lemishka I., Skrebtsov A., Ovchinnikov O. // Ukrainian Journal of Mechanical Engineering and Materials Science. 2017. Vol. 3, Issue 1. P. 73–80.

22. Ovchinnikov A. V., Ol'shaneckiy V. E., Dzhugan A. A. Primenenie nesfericheskih gidrirovanyh i degidrirovanyh poroshkov titana dlya polucheniya izdeliy v additivnyh tekhnologiyah // Vestnik dvigatelestroeniya. 2015. Issue 1. P. 114–117.

23. Petrik I. A., Ovchinnikov A. V., Seliverstov A. G. Razrabotka poroshkov titanovyh splavov dlya additivnyh tekhnologiy primeni-tel'no k detalyam GTD // Aviacionno-kosmicheskaya tekhnika i tekhnologiya. 2015. Issue 8. P. 11–16.

24. Broeke J., Perez J. M. M., Pascau J. Image Processing with ImageJ. 2nd ed. Packt Publishing, 2015. 256 p.

25. Gavrilova N. N., Nazarov V. V., Yarovaya O. V. Mikroskopicheskie metody opredeleniya razmerov chastic dispersnyh materialov: ucheb. pos. Moscow, 2012. 52 p.

26. Shanyavskiy A. A., Banov M. D., Zakharova T. P. Principles of physical mesomechanics of nanostructural fatigue of metals. I. Model of subsurface fatigue crack nucleation in VT3-1 titanium alloy // Physical Mesomechanics. 2010. Vol. 13, Issue 3-4. P. 133–142. doi: 10.1016/j.physme.2010.07.004

27. Model' Random Forest dlya klassifikacii, realizaciya na c#. URL: https://habr.com/post/215453

28. Random Forest: progulki po zimnemu lesu. URL: https://habr.com/post/320726

29. Information technologies of optimization of structures of the systems are on the basis of combinatorics methods / Oleg R., Yurii K., Oleksandr P., Bohdan B. // 2017 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT). 2017. doi: 10.1109/stc-csit.2017.8098776

30. Jordan M. I., Jacobs R. A. Hierarchical Mixtures of Experts and the EM Algorithm // Neural Computation. 1994. Vol. 6, Issue 2. P. 181–214. doi: 10.1162/neco.1994.6.2.181

31. Dronyuk I., Fedevych O., Poplavska Z. The generalized shift operator and non-harmonic signal analysis // 2017 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM). 2017. doi: 10.1109/cadsm.2017.7916092

32. Titanium Powder Metallurgy: Science, Technology and Applications / M. Qian, F. H. Froes (Eds.). Butterworth-Heinemann, 2015. 628 p. doi: 10.1016/c2013-0-13619-7

33. Titanium and Titanium Alloys: Fundamentals and Applications / C. Leyens, M. Peters (Eds.). Wiley-VCH Verlag GmbH & Co. KGaA, 2003. 532 p. doi: 10.1002/3527602119

34. Data Mining Fruitful and Fun. URL: https://orange.biolab.si

35. The Python Standard Library. URL: https://docs.python.org/3/library/index.html