nodules clearly shows the difference in the results. Classification accuracy for Naive Bayes is 99.5%, for Discriminant analysis - 87.7%, for Classification tree - 100%, for Nearest Neighbor - 100%.

Conclusion. The study considers the possibility of using fractal analysis for differential diagnostics of small pulmonary nodules on computer tomograms of patients with lung cancer. The methods of nonlinear mathematics showed significant differences in fractal dimension between pulmonary nodules of benign and malignant origin. Binary classifiers demonstrated the qualitative performance of the proposed research algorithm: decision trees (Classification tree) and k-of the nearest neighbors (Nearest Neighbor) distinguish between benign changes and malignant pathology by 100%.

Keywords: computed tomography, fractal analysis, lung nodules, lung cancer.

Introduction

Modern methods of X-ray computer tomography (CT) make it possible to detect small nodules (<10 mm) in the lung fields, which can be interpreted as metastatic in patients with malignant tumors. Pulmonary nodules <8 mm in size are considered small, nodules 3-5 mm in diameter are classified as very small, and the term "micronodules" is used for nodules <3 mm. In patients without malignant tumors, these "micronodules" should be considered benign until proven otherwise [1].

The radiographic features of benign and malignant nodules are remarkably similar, but several prognostic characteristics have been identified to aid in risk stratification. According to the Fleischner Society guidelines [2], these include nodule size, volume doubling time (VDT), margin characteristics, calcification, density, and location [3]. A significant increase in volume (>25%) with a volume doubling time (VDT) of less than 600 days between baseline and the last CT scan suggests nodule malignancy, with VDT ranging from 93 to 447 days [1].

Usually, in clinical practice, small, indeterminate nodules often require monitoring with serial CT scans of the chest for two years [2]. For these nodules, PET-CT with 18-FDG (Positron emission computer tomography with 18-fluorodeoxyglucose) is an imprecise method for distinguishing between benign and malignant lesions. Such patients lose a lot of time to confirm or deny metastatic spread. There is also a high risk of human error when

Chapter 10

The Method of Using Fractal Analysis for Metastatic Nodules Diagnostics on Computer Tomographic Images of Lungs

S. V. Romaniv[1]
Yu. B. Palaniza[1,2]
D. V. Vakulenko[1]
and I. Y. Galaychuk[1]

[1] I. Horbachevsky Ternopil National Medical University, Ternopil, Ukraine
[2] Ternopil Ivan Puluj National Technical University, Ternopil, Ukraine

Abstract

Purpose. To substantiate the use of method based on fractal analysis for differential diagnosis of small pulmonary nodules on computer tomograms (CT) of patients with lung cancer.

Materials and methods. The computer tomograms of 55 patients with lung cancer were analyzed retrospectively in dynamics. 195 lung nodules were selected on CT images, from which two cohorts of objects for fractal research were formed: "benign" (n = 120) and "malignant' (n = 75) nodules. Fractal analysis for each nodule was implemented in the Matlab environment by the method of counting boxes. 4 types of classifiers were used, their accuracy and error-rate were calculated.

Results. Fractal dimension along the x-axis (mathematical expectation from the dimension curve) amounts from 1.965 to 2.0, and along the y-axis (dispersion from the dimension curve) - from 0 to 0.025 for the group of "benign" nodules; fractal dimension along the x-axis was from 1.1 to 2.0, and along the y-axis - from 0 to 0.9 for the group of "malignant" nodules. The combined scattergram for both groups of

In: Horizons in Cancer Research. Volume 85
Editor: Hiroto S. Watanabe
ISBN: 979-8-88697-698-4

interpreting these images. Therefore, the search for non-invasive computer methods for the differential diagnosis of benign and malignant pulmonary nodules is becoming relevant.

In the last decade, CT techniques for detecting pulmonary nodules using artificial intelligence and machine learning have been developed [4]. Radiomics transfers imaging data into a high-dimensional feature space, using a large number of algorithms, and computational methodologies to transform medical images of tumors into quantitative descriptors that a computer extracts from the image and analyses them with the help of artificial intelligence.

Radiomic signs of malignant lung tumors differ depending on the morphological structure of the cancer, that is, CT images can differentiate the phenotypic difference of various histological variants of lung cancer [5]. Fractal analysis of medical images reveals physical characteristics and provides additional information about the complex structure of the image, where the features of the image are quantitatively obtained with the help of software, and the unevenness, heterogeneity of their surface texture are assessed. Some publications indicate that the fractal analysis of CT images of the lungs can be used to detect nodules for the differential diagnosis of lung tumors, as well as to distinguish their histological subtypes [6-9].

Considering the fact that malignant tumors are heterogeneous and inhomogeneous on CT images, and fractal analysis reflects the structural heterogeneity of the studied object, it would be appropriate to use it as an additional characteristic in the assessment of pulmonary nodules on CT.

The aim of the research. To substantiate the application of the technique based on fractal analysis for the differential diagnosis of small pulmonary nodules on computer tomography (CT) scans of lung cancer patients.

Research Materials and Methods

The retrospective research, which was conducted on the basis of Ternopil Regional Clinical Oncology Dispensary in the period from 2019 to 2020, included 55 patients with stage II-III-IV lung cancer (46 men, 9 women) aged 45-65 years (average 56,7 ± 4,7 years).

Criteria for the inclusion of patients in the research: morphologically confirmed lung cancer, the presence of additional pulmonary nodules from 2-3 to 10 mm in size on CT, in addition to the main tumor, at least two CT examinations in dynamics (after 6 and 12 months). A total of 195 small lung lesions were retrospectively selected, including non-solid nodules of the

"frosted glass" type, solid (dense, completely composed of solid components), and mixed-type nodules. According to the recommendations of the Fleischner Society [2] (2017), the dynamics of the size, volume and doubling time of the selected nodules were traced, on the basis of which two cohorts of objects for the fractal research were formed. Pulmonary nodules with stable CT characteristics formed the "benign" cohort (n = 120 nodules). The "malignant" cohort (n = 75 nodules) was formed from other nodules that underwent changes in CT dynamics, i.e., metastatic transformation.

CT examination of the chest cavity was performed on a Philips Brilliance 64 CT Scanner, standard protocols were used, and measurement was performed in phase with contrast and under the condition of the same slice thickness of 3 mm.

Figure 1 schematically shows the stages of the conducted research step by step. The lung tomogram scan with the nodule under study in DICOM format was opened in the Image-J software package [10] for image processing and calibrated using the CT Window Level plugin in the Lung Window (LUNG) in the same program. Next, the studied area (nodule) was selected using the "oval" tool. The nodule was completely absorbed into the area of the oval, avoiding lung tissue that could distort the measurements. A separate clean window (frame) with the size of 10 × 10 pixels 8 bits on a black background was created in advance, into which the selected nodule was copied and saved in jpeg format.

Description of the Methodology

Fractal analysis was implemented for each individual generated image in the Matlab environment [11]. The fractal dimension was calculated by the method of box counting [12] according to formula (1):

$$df = \lim_{\varepsilon \to 0} \frac{\log N(\varepsilon)}{\log (1/\varepsilon)}, \quad (1)$$

where df - fractal size of the nodule; ε - lateral length of the box; $N(\varepsilon)$ - the smallest number of contiguous and non-intersecting boxes of side length ε, required to completely cover an object (i.e., a nodule).

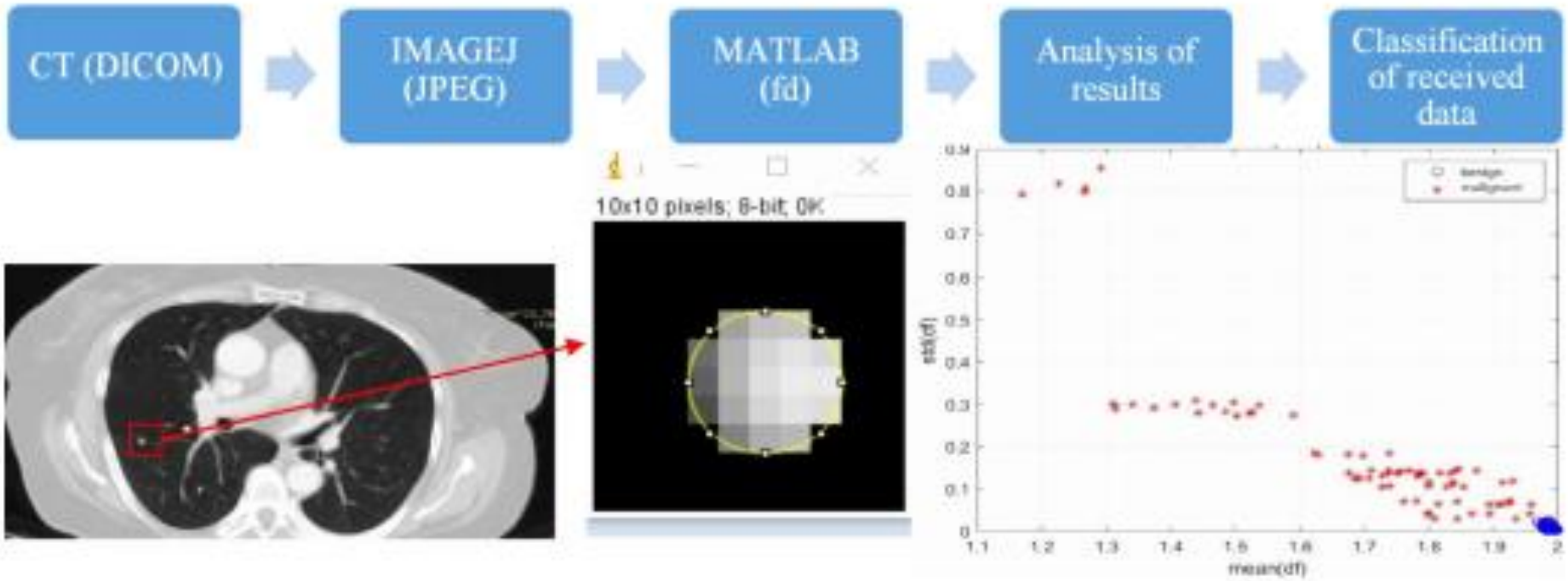

Figure 1. Workflow- diagram (upper part) of pulmonary nodule examination (CT - ImageJ - Matlab - analysis of results - classification of received data); (Bottom part): selection of a pulmonary nodule on a CT image - saving in jpeg format - fractal analysis of the nodules (scattergram).

Statistical processing of the research results was carried out in Matlab. Such indicators as mathematical expectation ***mean(df)*** (2) and standard deviation ***Std(df)*** (3) were calculated for each nodule:

$$\boldsymbol{mean\ (df)[x] = \int_a^b x f(x)\, dx} \tag{2}$$

where ***mean(df)*[x]** - continuous random variable with probability distribution density ***f(x)*** and focused on the interval ***(a,b)***.

$$\boldsymbol{Std(df) = \sqrt{\sum \frac{|x-\bar{x}|^2}{n}}} \tag{3}$$

where ***Std (df)*** - standard deviation indicator, x - average value of the sample (average fractal dimension), n - sample size (number of nodules).

In the Matlab software, 4 types of classifiers were used: Naive Bayes classifier, Discriminant analysis, Classification tree, Nearest Neighbor - to determine the probability of the fractal dimension of a nodule belonging to one of the classes of different classifiers. A matrix of inconsistencies was built as well as the error rate and classification accuracy were calculated.

Research Results

Graphs of the fractal analysis calculated by the box counting method for the "benign" and "malignant" nodule cohorts are shown in Figure 2.

For greater clarity, the graphs were transformed using the logarithmic formula (4):

$$\boldsymbol{df = -\mathbf{diff}(\mathbf{log}(\mathbf{n}))./\mathbf{diff}(\mathbf{log}(\mathbf{r}))} \tag{4}$$

where ***df*** - fractal dimension, ***n*** - number of boxes, ***r*** - size of the boxes.

After the transformation, the graph of the "benign" cohort looks less "complicated" than the graph of the "malignant" nodule cohort. The graph in Figure 3 (on the right) characterizes a more "complicated" structural heterogeneity typical for nodules of a malignant character.

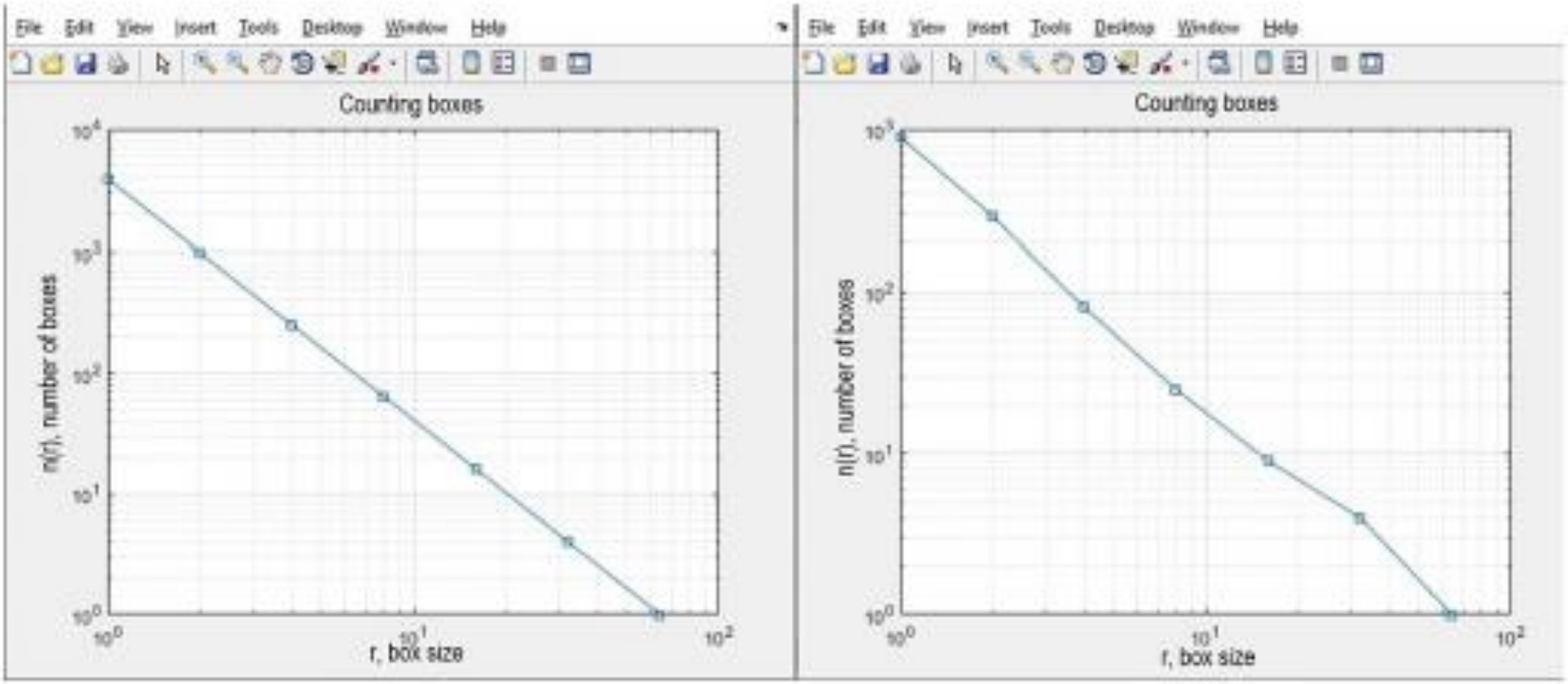

Figure 2. Graphs of the fractal dimension of pulmonary nodules, calculated by the box counting method*: on the left - the "benign" group, on the right - the "malignant" group.*

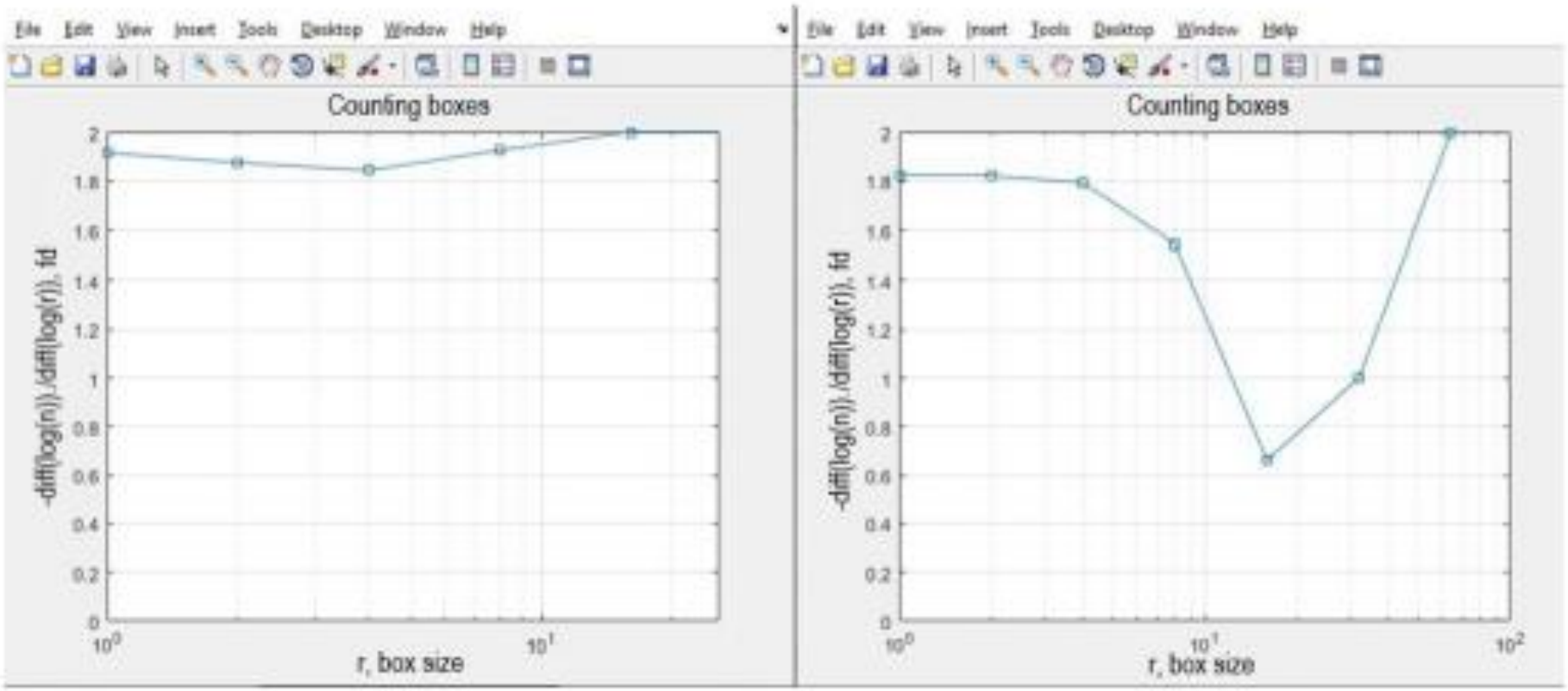

Figure 3. Graphs of the fractal dimension after logarithmic transformation: *on the left - "benign," on the right - "malignant."*

After calculating the indicators of mathematical expectation ***mean(df)*** formula (2) and standard deviation ***Std(df)*** formula (3), scattergrams were constructed for the "benign" and "malignant" nodule cohorts (Figures 4-5).

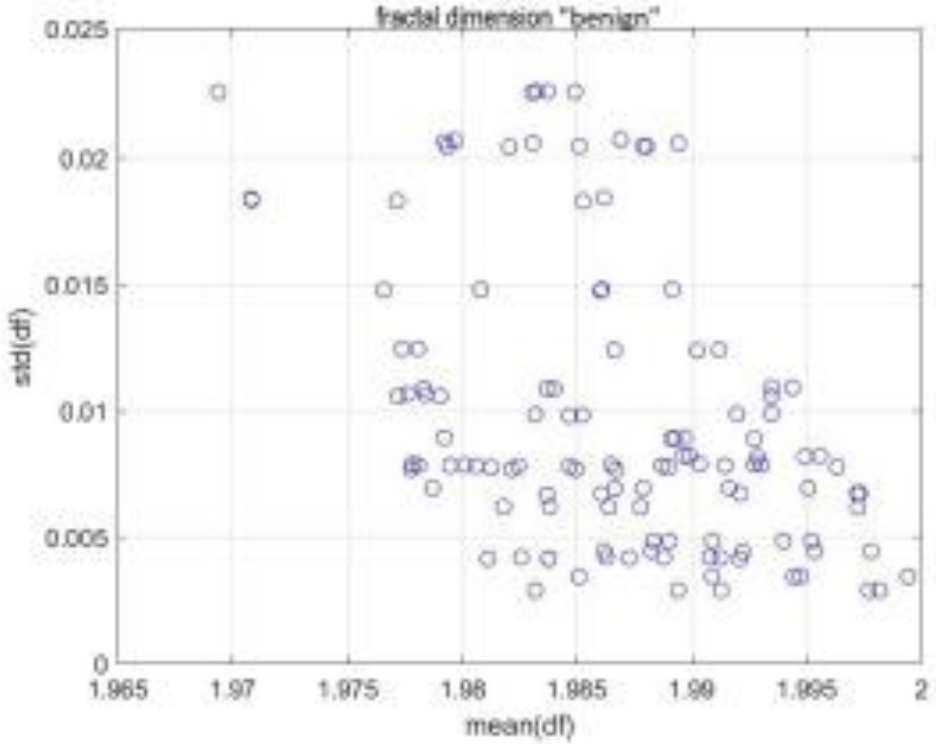

Figure 4. Fractal dimension scattergram of the "benign" nodules cohort.

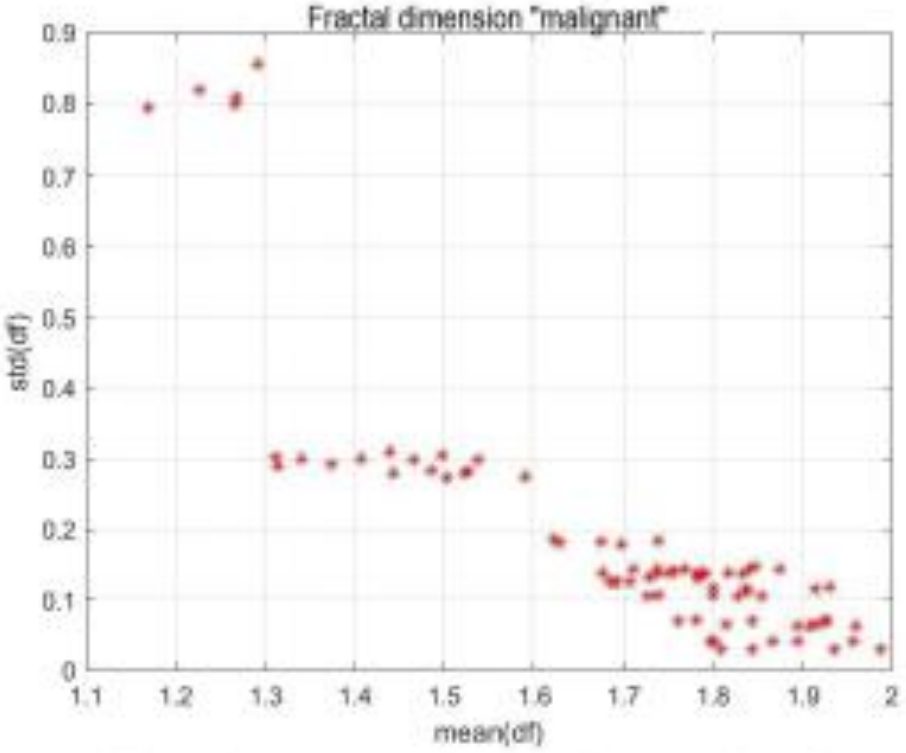

Figure 5. Fractal dimension scattergram of the "malignant" nodules cohort.

When combined into one coordinate system, the difference in the fractal dimension of the two cohorts of nodules is clearly visualized (Figure 6).

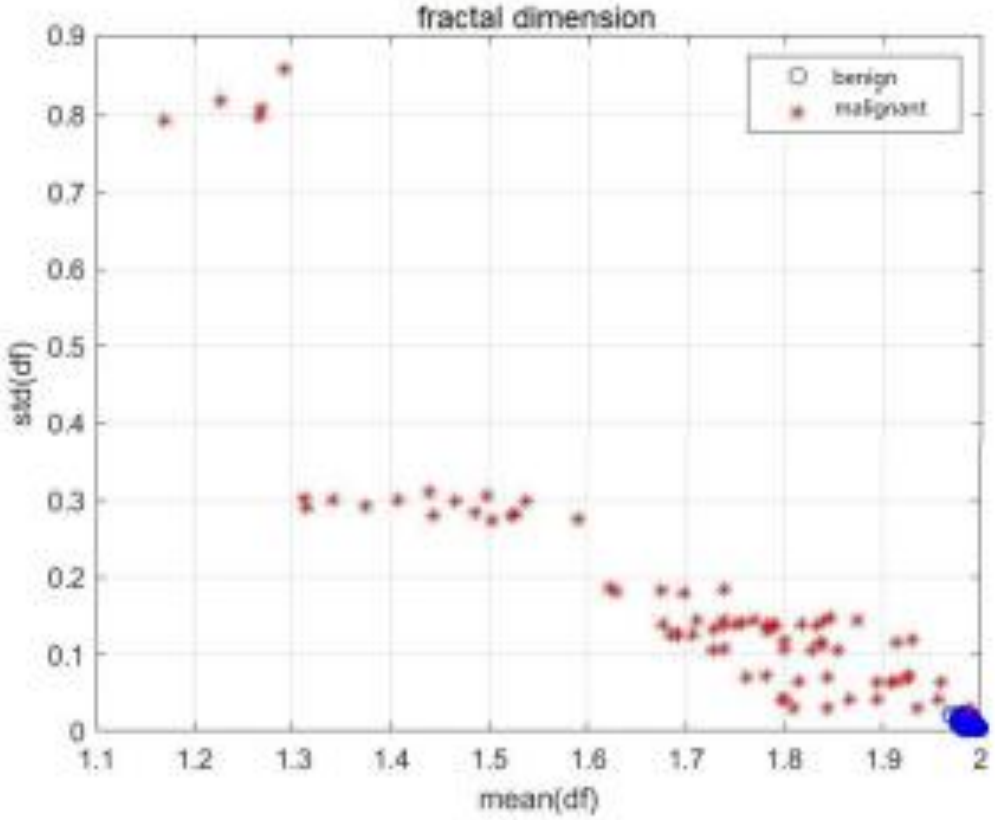

Figure 6. Pooled scatterplots of "benign" and "malignant" nodules cohorts: *asterisk markers represent placement of "malignant" nodules and ring markers - "benign" ones.*

In the feature space introduced by us for the "benign" cohort, the fractal dimension on the "*x*" axis (mathematical expectation from the dimension curve) is from 1.965 to 2.0; and on the "*y*" axis (dispersion from the dimension curve) - from 0 to 0.025; for the "malignant" cohort, the fractal dimension on the "*x*" axis is from 1.1 to 2.0; and on the "*y*" axis - from 0 to 0.9.

In order to check the fractal dimension of each individual nodule taken and its belonging to the corresponding cohort of "benign" or "malignant" ones, 4 types of feature classifiers were applied: Naive Bayes, Discriminant analysis, Classification tree, Nearest Neighbor.

The work of binary classifiers of 4 types is illustrated in Figures 7 and 8. In the lower corner of the coordinate space of all types of classifiers in Figure 7 data for the "benign" cohort are marked in light gray.

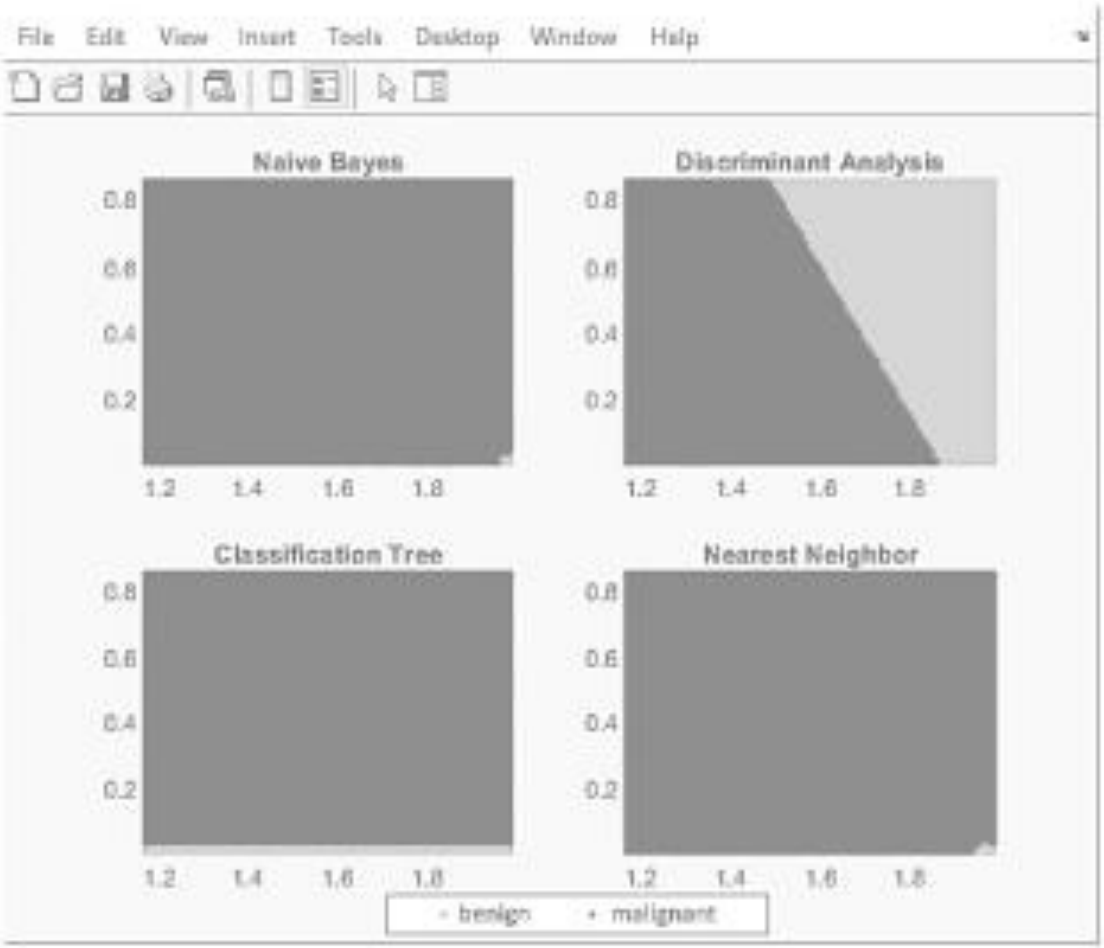

Figure 7. The work of binary classifiers. The dark gray colour characterizes the relation to the "malignant" cohort, and the light gray - to the "benign" cohort.

To check the work of the classifiers, a confusion matrix was built (Figure 8).

Class 0 includes nodules from the "benign" cohort (n = 120), and class 1 includes nodules from the "malignant" cohort (n = 75). The discrepancy matrix showed that Naive Bayes classifies one case (from the "malignant" cohort) to class 0 ("benign"), and Discriminant analysis classifies 24 cases of the "malignant" cohort to class 0 of the "benign" group, which is imprecise. And accordingly, the Classification tree and Nearest Neighbor correctly distributed the "benign" and "malignant" cohorts.

The error rate for Naive Bayes is 0.5%, for Discriminant analysis - 12.3%, for Classification tree - 0%, for Nearest Neighbor - 0%.

Classification accuracy for Naive Bayes is 99.5%, for Discriminant analysis - 87.7%, for Classification tree - 100%, and for Nearest Neighbor - 100%.

Figure 8. Application of classifiers (Naive Bayes, Discriminant analysis, Classification tree, Nearest Neighbor) to distinguish between "benign" and "malignant" cohorts.

After the obtained results, it was concluded that the classifier of the Classification tree and k-Nearest Neighbor distinguish nodules of "benign" and "malignant" cohorts by 100%, Naive Bayes distinguishes them by 99.5%, and Discriminant analysis - by 87.7%. This shows that conducted algorithm of fractal analysis distinguishes images in cohorts of benign and malignant nodules with a high degree of reliability.

Discussion

CT diagnosis of small lung formations up to 10 mm in size is complex; it always requires more effort and time to interpret the results. The goal of radiological assessment of small pulmonary nodules is to differentiate benign from malignant formations as accurately and non-invasively as possible.

Follow-up CT scans every three months, PET/CT with FDG, or nodule biopsy are the next diagnostic steps [2]. In most cases in the presence of nodules >8 mm on LDCT (low-dose CT), PET/CT with FDG is helpful in making a further decision regarding biopsy of the nodule, but in 28% of cases additional studies were required. However, for nodules <8 mm, PET/CT is a less sensitive method [13].

For these types of nodules, there is a textural analysis of medical images that allows to quantitatively evaluate the characteristics of the examined object, regardless of size, to provide a radiologist with an additional assessment and to help in making the right decision. Texture analysis also includes fractal analysis.

Fractal analysis has already been previously applied for the diagnosis of malignant and benign lung pathologies, while the parameters of fractal analysis on CT (m-FD), on PET/CT (d-FD) and SUVmax on PET/CT were compared. The optimal limit values for differentiating malignant from benign pulmonary nodules were 1.183, 4.24, and 0.0267 for m-FD, SUVmax, and d-FD, respectively. The diagnostic accuracy of SUVmax (68.5%) and d-FD (77.8%) on PET was better than that of m-FD on CT (64.8%). Combining m-FD with SUVmax or d-FD increased accuracy to 92.6% and 94.4%. The results of this research indicate that SUVmax and d-FD on PET/CT images provide different information about nodules and are equally useful for differential diagnosis [14].

Another research proved the possibility of a radiomic model in distinguishing histological types of tumors based on quantitative characteristics. The machine learning model showed the highest spaces under the curves AUC 0.92, 0.84 and 0.88 for distant metastases, regional metastases, and primary lung cancer (adenocarcinoma, squamous cell carcinoma), respectively. One of the quantitative features in this study was the fractal dimension [15-17].

The method of fractal analysis, presented in our research, can help radiologists in making the right decision regarding the assessment of the nature of a pulmonary nodule, and provide new information characteristics about the object under study. Thus, the visual assessment of the pathological formation is complemented by its mathematical processing using the methods of nonlinear mathematics, followed by the analysis of the converted digital image into a graphic format and the corresponding interpretation.

Conclusion

1. The research considered the possibility of using fractal analysis for the differential diagnosis of small pulmonary nodules on computer tomograms of patients with lung cancer.
2. Significant differences in the fractal dimension between benign and malignant pulmonary nodules are shown using the methods of non-linear mathematics.
3. Binary classifiers demonstrated the qualitative work of the proposed research algorithm: Classification tree and k-Nearest Neighbor distinguish between benign changes and malignant pathology by 100%.

Prospects for Future Research

The following studies will be aimed at improving the existing algorithm in order to increase the sensitivity and cover a larger cohort of studied objects.

Acknowledgments

The study was carried out within the framework of the scientific research work of I. Ya. Horbachevskyi Ternopil National Medical University under the Ministry of Health of Ukraine: "Optimization of radiological diagnostics in patients with lymphoma and lung cancer," state registration number: 0122U000161.

Project 2021.01/0136 "Development of a cloud platform for patient-centered telerehabilitation of oncology patients based on mathematical modelling," 2022-2023, "Science for the safety of man and society" contest of the National Research Fund of Ukraine.

References

[1] Borghesi, A., Michelini, S., Nocivelli, G., Silva, M., Scrimieri, A., Pezzotti, S., Maroldi, R., & Farina, D. (2019). Solid indeterminate pulmonary nodules less than or equal to 250 mm3: Application of the updated fleischner society guidelines in

clinical practice. *Radiology Research and Practice*, *2019*, 1-7. https://doi.org/10.1155/2019/7218258.

[2] MacMahon, H., Naidich, D. P., Goo, J. M., Lee, K. S., Leung, A. N. C., Mayo, J. R., Mehta, A. C., Ohno, Y., Powell, C. A., Prokop, M., Rubin, G. D., Schaefer-Prokop, C. M., Travis, W. D., Van Schil, P. E., & Bankier, A. A. (2017). Guidelines for management of incidental pulmonary nodules detected on CT images: From the fleischner society 2017. *Radiology*, 284(1), 228-243. https://doi.org/10.1148/radiol.2017161659.

[3] Cruickshank, A., Stieler, G., & Ameer, F. (2019). Evaluation of the solitary pulmonary nodule. *Internal Medicine Journal*, 49(3),306-315. https://doi.org/10.1111/imj.14219.

[4] Halder, A., Dey, D., & Sadhu, A. K. (2020). Lung nodule detection from feature engineering to deep learning in thoracic CT images: A comprehensive review. *Journal of Digital Imaging*, 33(3), 655-677. https://doi.org/10.1007/s10278-020-00320-6.

[5] Zhu, X., Dong, D., Chen, Z., Fang, M., Zhang, L., Song, J., Yu, D., Zang, Y., Liu, Z., Shi, J., & Tian, J. (2018). Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. *European Radiology*, 28(7), 2772-2778. https://doi.org/10.1007/s00330-017-5221-1.

[6] Linning, E., Lu Lin, Li, L., Yang H., Schwartz. L., H., Zhao. B. (2019). Radiomics for classification of lung cancer histological subtypes based on nonenhanced computed tomography. *Academic Radiology*, 26(9), 1245-1252. https://doi.org/10.1016/j.acra.2018.10.013.

[7] Naik, A., Edla, D. R., & Dharavath, R. (2021). Prediction of malignancy in lung nodules using combination of deep, fractal, and gray-level co-occurrence matrix features. *Big Data*. 9(6), 480-498. https://doi.org/10.1089/big.2020.0190.

[8] Lee, C. H., & Jwo, J. S. (2020). Automatic segmentation for pulmonary nodules in CT images based on multifractal analysis. *IET Image Processing*, 14(7), 1347-1353. https://doi.org/10.1049/iet-ipr.2019.0884.

[9] Marceniuk, V. P., Halaychuk, I. Y., Zhulkevich, I. V. Nazarkevich, H. P. (2013). Compiuterna programa [Complete the program] "Informaciyna systema dyferenciynoii diagnostyky dribnych vnutrishniolehenevych utvoriv na compiuternych tomohramach (LUNG-CT) [*The information system of differential diagnosis of trifles was created on a computer tomograph (LUNG-CT)*]." *Avtorske pravo na tvir* №49686 [in Ukrainian].

[10] Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), 671-675. https://doi.org/10.1038/nmeth.2089.

[11] MathWorks, I. (2009). *MATLAB 7: Getting Started Guide*. The MathWorks, Inc.

[12] Jundong Yan, Yuanyuan Sun, Shanshan Cai & Xiaopeng Hu. (2016). An improved box-counting method to estimate fractal dimension of images. *Journal of Applied Analysis & Computation*, 6(4), 1114-1125. https://doi.org/10.11948/2016073.

[13] Hadique, S., Jain, P., Hadi, Y., Baig, A., & Parker, J. E. (2020). Utility of FDG PET/CT for assessment of lung nodules identified during low dose computed

tomography screening. *BMC Medical Imaging*, 20(1). https://doi.org/10.1186/s12880-020-00469-0.

[14] Miwa, K., Inubushi, M., Wagatsuma, K., Nagao, M., Murata, T., Koyama, M., Koizumi, M., & Sasaki, M. (2014). FDG uptake heterogeneity evaluated by fractal analysis improves the differential diagnosis of pulmonary nodules. *European Journal of Radiology*, 83(4), 715-719. https://doi.org/10.1016/j.ejrad.2013.12.020.

[15] Ferreira-Junior, J. R., Koenigkam-Santos, M., Magalhães Tenório, A. P., Faleiros, M. C., Garcia Cipriano, F. E., Fabro, A. T., Näppi, J., Yoshida, H., & de Azevedo-Marques, P. M. (2019). CT-based radiomics for prediction of histologic subtype and metastatic disease in primary malignant lung neoplasms. *International Journal of Computer Assisted Radiology and Surgery*, 15(1), 163-172. https://doi.org/10.1007/s11548-019-02093-y.

[16] Martseniuk, V., Lupenko, S., Semenets, A., Vakulenko, D., Kravets, N., & Klymuk, N. (2021) On Data Mining Technique for Differentiation Condition of Football Players Using of Arterial Oscillography. *Proceedings of 11th International Conference on Advanced Computer Information Technologies (ACIT)*, 662-665. https://doi.org/10.1109/ACIT52158.2021.9548644.

[17] Martsenyuk, V.P., Vakulenko, D.V., Hryshchuk, L.A., Vakulenko, L.O., Kravets, N.O., Klymuk, N.Y. (2022) On the Development of Directed Acyclic Graphs in Differential Diagnostics of Pulmonary Diseases with the Help of Arterial Oscillogram Assessment. *Mechanisms and Machine Science* 107:157-173. https://doi.org/10.1007/978-3-030-76787-7_8

Biographical Sketches

Romaniv S. V.

PhD student of the Department of Oncology, Radiodiagnosis, Radiotherapy and Radiation Medicine, Ivan Horbachevsky Ternopil National Medical University of the Ministry of Health of Ukraine;. ORCID ID 0000-0002-2802-1087.

Palaniza Y. B.

Candidate of Engineering Sciences (Ph. D.), Senior Lecturer of Radiotechnical systems Dep., Ternopil Ivan Puluj National Technical University, Ukraine. Candidate of Engineering Sciences (Ph. D.), Assistant of Medical Informatics Dep., I. Horbachevsky Ternopil National Medical University, Ukraine. ORCID ID 0000-0002-8710-953X.

Vakulenko D. V., DSc

Professor, Head of Medical Informatics Dep., Ivan Horbachevsky Ternopil National Medical University of the Ministry of Health of Ukraine. ORCID ID: 0000-0001-5062-9278.

Galaychuk I. Y., MD, DSc

Professor, Head of the Department of Oncology, Radiodiagnosis, Radiotherapy and Radiation Medicine, Ivan Horbachevsky Ternopil National Medical University of the Ministry of Health of Ukraine. ORCID ID: 0000-0003-1863-4709.