Recognition of Diffuse Hepatic Steatosis

Danil Yevdokimov, Viacheslav Gorikhovskii
Saint Petersburg State University
Saint Petersburg, Russian Federation
st084331@student.spbu.ru, v.gorikhovskii@spbu.ru

***Abstract*—Since the COVID-19 pandemic, chest computed tomography has become a common practice, with the liver and spleen being studied in addition to the lungs and heart. Unfortunately, radiologists often do not have the time to assess the density of the liver's structure, as their attention is primarily on the chest organs such as the lungs, heart, and blood vessels in the thorax. To address this issue, a solution could be the development of a tool that uses machine learning technologies and statistical methods to identify organs and their relationships to diagnose diffuse hepatic steatosis on chest computed tomography scans. Based on an open dataset, various methods of statistical and regression analysis were used, and diagnostic tests were obtained with more than 95% accuracy.**

I. Introduction

Hepatic steatosis is a frequent observation during imaging studies that may be indicative of chronic liver ailments, with non-alcoholic fatty liver disease being the most prevalent. It refers to an atypical buildup (up to 5%) of triglycerides in the hepatocytes [1]. This constitutes a major public health issue, with estimates suggesting that the prevalence of hepatic steatosis among the population in Western countries could be as high as 30% [2].

Liver biopsy is the established method for diagnosing non-alcoholic fatty liver disease, as it provides histolog-ical evaluation that can reveal features not detectable on imaging, such as specific patterns of inflammation and hepatocyte damage [3]. However, the procedure is associated with high costs, sampling error, and potentially life-threatening complications. In contrast, non-invasive imaging techniques, including ultrasonography, computed tomography (CT), magnetic resonance spectroscopy, and magnetic resonance imaging (MRI), are increasingly being used to assess important disease markers such as hepatic steatosis and advanced liver fibrosis.

Ultrasound is a safe, affordable study, but has several disadvantages for detecting and assessing steatosis, for example, in patients with a large body mass index, the steatosis estimate may be overestimated due to beam attenuation when passing through abdominal and visceral fat, and not liver fat, liver echogenicity may be distorted by fibrosis, inflammation and other signs of chronic liver disease, fibrosis and fat may superficially resemble each other [4]. In addition, conventional ultrasound is operator dependent, resulting in varying results and reproducibility. MRI is considered the most sensitive and specific method for assessing steatosis. Unlike ultrasound and computed tomography, which measure steatosis indirectly, MRI measures the signal intensity (brightness) of protons at different resonant frequencies [5]. MRI uses the difference in proton resonance frequencies between water and triglycerides. However, the MRI method has a number of limitations in use, is relatively expensive, and also cannot easily distinguish between a fat fraction of more than 50% and a fat fraction of less than 50%.

Therefore, computed tomography without contrast enhancement was proposed as an alternative to liver biopsy. The main advantages of computed tomography for assessing steatosis are relatively fast data acquisition, ease of execution, simple analysis and quantitative results. Since the COVID-19 pandemic, computed tomography scans of the chest have become a standard study that also includes a focus on the liver and spleen. A normal liver parenchyma on unenhanced computed tomography scans appears brighter than the spleen, with a density of around 60 HU [6].

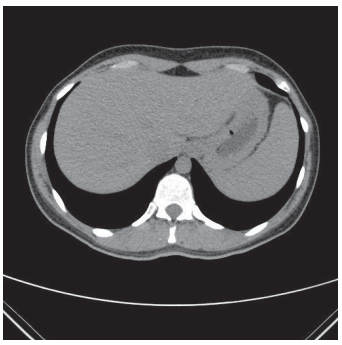

Fig. 1. Computed tomography scan of the liver without steato-sis

However, in cases of steatosis, the liver tissue becomes hypodense compared to the surrounding spleen.

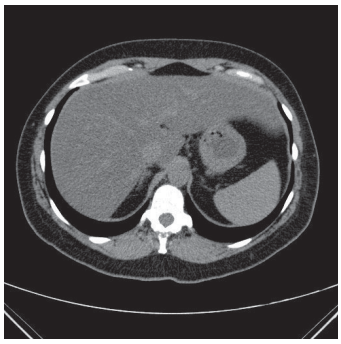

Fig. 2. computed tomography scan of the liver with steatosis

Despite this, radiologists may not have enough time to evaluate liver structure due to the emphasis on chest organs such as the lungs, heart, and thoracic blood vessels. To address this issue, a solution could be the development of a tool that uses machine learning technologies and statistical methods to recognize organs and identify correlations between available data and the presence of diffuse hepatic steatosis on chest computed tomography scans. This tool is intended to be purely advisory. If it reports the possible presence of hepatic steatosis, then the patient is advised to undergo some specialized test for diseases. For example, such non-invasive methods as SteatoTest [7] or FibroScan [8].

II. Overview

A. Overview of existing solutions

In 2020, the Scientific and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow City Health Department published data on the detection of liver steatosis in 30,000 patients who underwent computed tomography [9]. For verification, the program of automatic liver densitometry CTLiverExam was used, which allows assessing the density of the liver even in conditions of its partial capture in the scanning area [10]. The result of batch processing of images was given in the form of a table with data on the average density in the segmented area, standard deviation, degree of heterogeneity. Segmentation was considered incorrect in cases where the selected area included the stomach, heart, ribs, right-sided hydrothorax, and less than 90% of the liver. Despite the results obtained, this method has not been introduced into practical healthcare.

A group of authors from the USA and Germany in 2021 published a paper on the detection of liver steatosis according to chest computed tomography. The study included 15,000 patients. A deep learning program was created, liver fat quantification was measured in Hounsfield units by three independent methods: 1. Average attenuation of the region of interest in three cross-sectional images at different levels of the liver; 2. Average attenuation on images manually selected by the doctor; 3. Average attenuation of volumetric segmentation based on deep learning. The disadvantage of the method is the inability to interpret images with low image quality [11]. It should be noted that every year you can find publications devoted to the detection of liver steatosis according to native CT, it should be noted that in general the detection algorithm is approximately the same for everyone, and is based on Hounsfield units, however, this method does not work when analyzing enhanced CT and CT with contrast [12].

TABLE I. Technology comparsion

Name	Free	Pretrained model for liver
Visiopharm	-	+
Halo AI	-	+
Aiforia	-	+
MONAI-Lable	+	-
H-AI-L	+	-
QuickAnnotator	+	-
livermask	+	+

B. Technology overview

It was decided to use the Python programming language for data collection and analysis, since this language was also used in other works related to the analysis of CT images and segmentation of organs on them [11], [13]–[15]. To obtain data on the brightness of the liver, it is necessary to segment it on CT. There are several popular tools for this purpose:

1) Visiopharm [16];
2) Halo AI [17];
3) Aiforia [18];
4) MONAI-Lable [19];
5) H-AI-L [20]
6) QuickAnnotator [21];
7) Library livermask [22]

Of all the tools presented above, livermask was chosen, since livermask is a library for automatic segmentation of the liver parenchyma on CT using deep learning, it has the following advantages:

1) Open source;
2) There is a pre-trained model for liver segmentation;
3) Regular updates;
4) Easy to use. To start segmentation, you need to call one function.

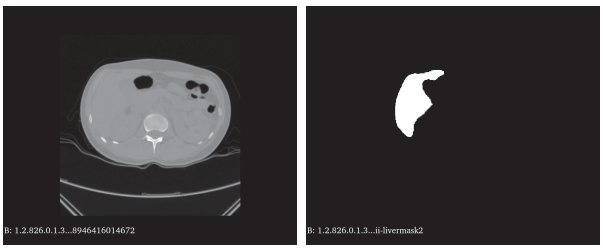

Fig. 3. An example of how livermask works

The sklearn [23] library is used to obtain the score to separate the data into a training sample and tests, for

linear and polynomial regression, since the main purpose library is classification and regression. In addition, this library was also used in other works related to the analysis of CT images and segmentation of organs on them [11], [14].
The nibabel [24]library is used to obtain a three-dimensional array of pixel brightness in a CT study. This package provides read +/- write access to several common medical and neuroimaging image file formats, including: GIFTI, NIfTI1, NIfTI2, CIFTI-2, MINC1, MINC2, AFNI BRIK/HEAD, MGH and ECAT, and Philips PAR/ REC.
To calculate the Spearman correlation coefficient, the scipy [25] library is used. SciPy is an open source library for performing scientific and engineering calculations.
For operations with matrices, the numpy library [26] is used. NumPy is an open source library for working with multidimensional arrays, including matrices.
For finding various averages, working with csv files, getting the current time, creating folders, getting a list of files in a folder, mathematical calculations, getting combinations, standard Python libraries such as statistics, datetime, os, glob, math and itertools are used.

C. Review of mathematical methods

Correlation analysis is used to search for the most significant features.
Statistical criteria and regression analysis are used to build diagnostic tests. This paper uses two criteria:

1) the most powerful criterion to find the point relative to which the steatosis result will be given;
2) fuzzy criterion is a function $\delta(\mathbf{X})$ that takes values in the interval $[0, 1]$ depending on the sample values. It is used to calculate the probability of having steatosis based on statistical data on the presence of steatosis in patients with a higher and lower liver brightness value.

Three metrics are used to evaluate diagnostic tests:

1) coefficient of determination (R^2 score). R^2 score is needed to evaluate the accuracy of tests with the same number of factors;
2) logarithmic loss (*log* loss). *log* loss is a metric for evaluating a binary classification model, in which the further the prediction probability is from its true value, the higher the penalty;
3) F-metric ($F1$ score) only for constructing the most powerful criterion, since $F1$ score is well suited for a binary result.

III. Background

This work was inspired by the Kaggle [27] to recognize diffuse hepatic steatosis on chest CT. This competition also provided a dataset consisting of 152 patients with chest CT and 112 patients with abdominal or abdominal and chest CT. The training set contains 226 patients, of which 114 were patients with chest CT.
I would also like to note that Ekaterina Zaslavskaya also made a great contribution to this work. She was our medical consultant, and she also manually marked the liver on computed tomography scans to compare them with the results that the computed produces.

IV. Implementation description

A. Data preparation

Since some patients have more than one study, it is necessary to convert the available data on the presence of steatosis into a format in which the minimum unit of division is not the patient, but a single study. This increases the training and test data from 226 to 301 and resolves the issue of which study should be considered correlated with steatosis data.
Then this data is divided into a training sample and tests using the sklearn library. It stratifies the data into two samples in the desired proportions. In my chosen case, the training sample consists of 70% of the available data, the test sample consists of 30% [28], [29]. There are 30 training splits and 30 test splits in total. Then the liver masks of each study are obtained using the livermask library.
Masks are saved in NifTi format.
After that, the collection of data on the brightness of the liver and the entire NifTi file begins. The brightness of a single point is the element value of the 3D pixel array of a single study.
About the liver, the following brightness data is collected:

1) Brightness value at a random point in the liver;
2) Arithmetic mean of brightness, median brightness, brightness mode, brightness quantiles of 100 random points in the liver;
3) Arithmetic mean of brightness, median brightness, brightness mode, brightness quantiles of the points of one, two and three random area of the liver;
4) Arithmetic mean of brightness, median brightness, brightness mode, brightness quantiles of the points of the entire liver.

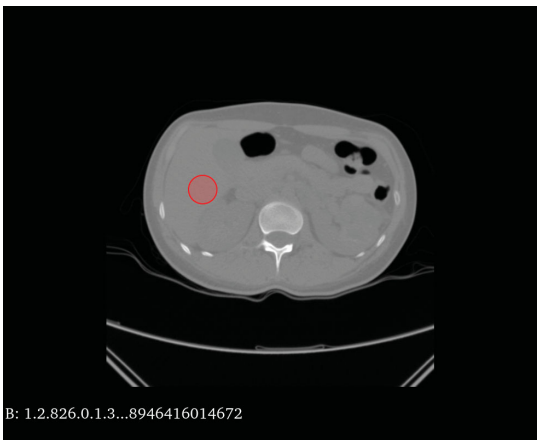

Fig. 4. An example of selecting a random area of the liver

TABLE II. PART OF THE DATA ON PATIENT NUMBER 3 WITH STEATOSIS AND PATIENT NUMBER 1 WITHOUT STEATOSIS

Value	Steatosis	No steatosis
Random point brightness	9.0	58.0
Random points mean	14.44	50.58
Random points median	12.0	51.0
Random points mode	21.0	60.0
Random points 1st quan.	-2.75	39.0
Random points 3rd quan.	36.0	60.0
One area mean	17.88	53.28
One area median	17.0	54.0
One area mode	15.0	56.0
One area 1st quan.	1.0	44.0
One area 3rd quan.	34.0	63.0
Whole liver mean	20.14	51.59
Whole liver median	20.0	52.0
Whole liver mode	18.0	54.0
Whole liver 1st quan.	3.0	43.0
Whole liver 3rd quan.	37.0	61.0

The following data is collected about the entire study:

1) The arithmetic mean of the brightness of the points of the whole study;
2) Median brightness of the points of the whole study;
3) Brightness mode of points of the whole study;
4) Brightness quantiles of the points of the whole study.

Sets of values are converted to numbers using the statistics library, which is equipped with a large number of functions to obtain statistically significant data. After that, the resulting lists of dictionaries are saved in the corresponding csv files.
Further, to obtain information about the significance of different brightness values, it is necessary to obtain the dependence of the available data on the value of the presence of steatosis. This can be done by finding the modulus of the Spearman rank correlation coefficient using the scipy libraries, which have functions for calculating the Spearman correlation coefficient, and numpy, which is necessary to write your own functions for calculating the multiple correlation coefficient. It is necessary to consider rank correlation, since the relationship between binary and continuous data is being sought. It is the modules that need to be looked for, since in order to find out the relationship between the set of brightness values and the presence of steatosis, it is necessary to calculate the coefficient of multiple correlations, whose sign is not possible to find out due to the properties of the square root [30].
The resulting correlation coefficients are stored in a list of dictionaries, and the list is converted to a csv file.

B. Experimental study

1) Experimental conditions: Spearman's multiple correlation coefficient is considered to determine the significance of each set of different brightness values for different parts of the liver. It is necessary to consider rank correlation, since the relationship between binary and continuous data is being sought.

2) Research questions: The following hypotheses were formulated:

1) the larger the selected part of the liver, the greater the correlation coefficient;
2) the greater the number of different brightness values, the greater the correlation coefficient.

3) Results of correlation analysis: During the collection of correlation coefficients between the brightness values and the presence of steatosis, it was found that the larger the liver area is taken, the more the studied values correlate. For the same reason, the smallest correlation coefficient was obtained for combinations with the value of the brightness of a random point.

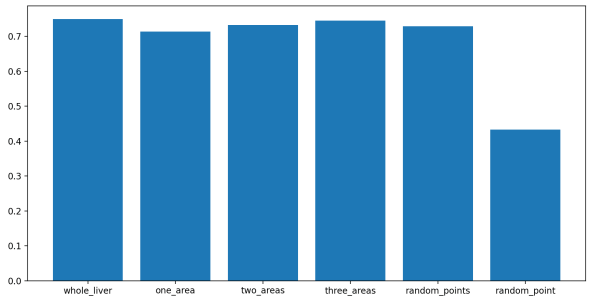

Fig. 5. The average value of the correlation coe icient for different parts of the liver

Excluding the case with a random point from the analysis, since it has only one kind of brightness value, it can be seen that the correlation coefficient increases with the increase in the number of different brightness values of different parts of the liver and the entire study.

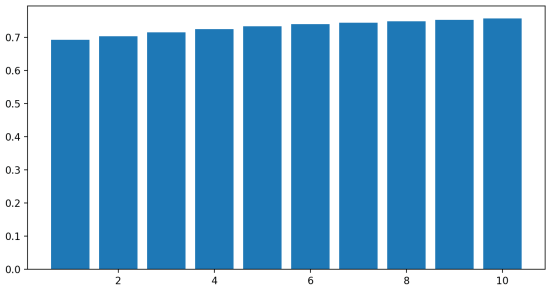

Fig. 6. The average value of the correlation coe icient over the number of different brightness values

From the data obtained, we can conclude that to obtain the best result, you should take as many liver points as possible and calculate as many different brightness values from them.

C. Predictive Models

Based on the results of the correlation analysis and the data obtained, the following models for the diagnosis of steatosis were implemented.

1) Construction of the most powerful criterion: From the data on the brightness of the parts of the liver from the training sample, we obtain the maximum and minimum values. Then the average value between them is taken. This point is necessary to determine the presence of steatosis. If the input is given a value greater than it, then there is no steatosis, otherwise there is. But there is no evidence that this particular point will give the minimum logarithmic loss. Therefore, relative to this point, the leftmost boundary is searched first, minimizing score. Then the right one. As a result, the point with the minimum score is selected from the two of them. And already relative to this point in the tests, a verdict will be issued about the presence or absence of steatosis. The test results are entered into the list of dictionaries. After that the list is saved as a csv file.
There is also a similar model that maximizes the f1 score.

2) Construction of a fuzzy criterion: Data on the brightness of different parts of the liver from the training sample are divided into 2 lists based on the presence of steatosis. Then the minimum is taken from the list of those without steatosis and the maximum from the list of those with steatosis. This is the minimum and maximum of the intersection of the lists. Then we assign the following scheme to the incoming value:

1) If the incoming brightness value is greater than the maximum of the intersection, then there is no steatosis;
2) If the incoming brightness value is less than the minimum of the intersection, then there is steatosis;
3) If the input value from the test fell into the intersection, then the ratio of the number of those with steatosis with a brightness value greater than the input to the size of the entire intersection will be returned.

The test results are entered into the list of dictionaries. The list is then saved as a csv file.

3) Linear regression: Linear regression [31] is performed on all combinations of brightness values of various parts of the liver and the entire study. Then they are tested. The test results are entered into the list of dictionaries. The list is then saved as a csv file.

4) Polynomial regression: Polynomial regression [32] of the second and third degree is carried out on all combinations of brightness values of various parts of the liver and the entire study. Further testing and recording the results obtained is similar to linear regression. At the end of data collection, the results are written to a csv file.

V. DIAGNOSTIC TEST RESULTS

During the diagnostic tests, the following results were obtained:

1) The most powerful criterion was tested with a best f1 score of 0.98 and a log loss of 1.14 when trained on the median brightness values of the whole liver. From the obtained data on the most powerful criterion, it can be concluded that in order to obtain the best result, it is worth using the median value of the brightness of the entire liver to find the boundary point, relative to which a verdict of the presence or absence of steatosis is issued;

2) The fuzzy criterion received the best R^2 score equal to 0.8 during testing, while the logarithmic loss is 0.5 when learning on the modes of the brightness values of the entire liver.

3) The linear regression model received the best R^2 score in testing equal to 0.61, while the logarithmic loss is 0.27 when trained on the arithmetic means, modes, first and third quantiles of the brightness values of the entire study and the arithmetic means, medians, first and third quantiles brightness values of three random areas of the liver.

4) The second-degree polynomial regression model obtained a best R^2 score of 0.76 on testing, with a log loss of 0.17 when trained on the modes of the luminance values of the entire study and the mean, median, and modes of the luminance values of three random liver regions.

5) The third-degree polynomial regression model obtained a best R^2 score of 0.81 when tested, with a log loss of 0.13 when trained on the third quantile of the whole study luminance value and all the whole liver luminance values.

TABLE III. CONFUSION MATRIX OBTAINED WHEN PASSING THE 20TH TEST USING THE MOST POWER-FUL CRITERION OF THE MEDIAN BRIGHTNESS OF THE WHOLE LIVER TRAINED ON THE 30TH TRAINING SET

	Predicted Positive	Predicted Negative
Positive	55	9
Negative	1	26

	place	type	min_log_loss	max_f1_score
0	whole_liver	mean	3.80	0.98
1	whole_liver	median	1.14	0.98
2	whole_liver	mode	1.90	0.96
3	whole_liver	1	3.42	0.92
4	whole_liver	3	3.04	0.95
5	one_area	mean	6.45	0.95
6	one_area	median	1.90	0.96
7	one_area	mode	2.28	0.95
8	one_area	1	4.18	0.91
9	one_area	3	2.28	0.96
10	two_areas	mean	2.66	0.95
11	two_areas	median	1.90	0.96
12	two_areas	mode	2.66	0.96
13	two_areas	1	3.42	0.92
14	two_areas	3	2.28	0.96
15	three_areas	mean	2.66	0.95
16	three_areas	median	2.28	0.96
17	three_areas	mode	2.66	0.94
18	three_areas	1	3.42	0.92
19	three_areas	3	2.66	0.95
20	random_points	mean	4.55	0.90
21	random_points	median	3.04	0.95
22	random_points	mode	6.07	0.93
23	random_points	1	4.18	0.90
24	random_points	3	2.66	0.94

Fig. 7. A table where each pair, liver part and brightness value, corresponds to the minimum logarithmic loss and the highest f1 score obtained when testing the most powerful criterion

TABLE IV. CONFUSION MATRIX OBTAINED WHEN PASSING THE 29TH TEST USING THE MOST POWER-FUL CRITERION OF THE MEAN BRIGHTNESS OF THE WHOLE LIVER TRAINED ON THE 11TH TRAINING SET

	Predicted Positive	Predicted Negative
Positive	55	10
Negative	1	25

A. Discussion of results

It can be noted that the results of diagnostic tests of polynomial regression do not agree with the conclusions of the correlation analysis. Increasing the number of variables does not improve the score. This may be due to the inadequacy of the sample size, since it is believed that the number of observations, according to various estimates, should be at least 10 times greater than the number of factors. If the number of observations exceeds the number of factors by dozens of times, the law of large numbers comes into play, which ensures the mutual cancellation of random fluctuations [33]. There are only 301 observations in the existing sample, so perhaps the number of observations should exceed the number of factors by more than 30 times. To test this assumption, it is necessary to increase the sample.

In addition, I would like to note that all the models listed in the work were trained and tested on standard computed tomography scans. The behavior of each model remains unknown if enhanced CT or contrast-enhanced CT is input. To solve this problem, new models for diagnosing steatosis trained on various types of CT are required.

CONCLUSION

In the course of the work, the following tasks were completed:

- the signs that give the greatest correlation with the presence or absence of steatosis were identified;
- a mathematical model was built that predicts the presence of liver steatosis according to computed tomography data;
- implemented a program code that allows you to identify sufficient data to determine liver steatosis according to computed tomography;
- implemented diagnostic tests built on the basis of a mathematical model.
- code tested with data received from Kaggle.

The following is also planned for the future:

- collection of data on the brightness of the spleen;
- identification of signs that give the highest correlation with the presence or absence of steatosis on new data;
- implementation of new diagnostic tests;
- development of an application using the most efficient models;
- testing and implementation of the application.

Application prototype on GitHub that can report the possibility of liver steatosis for the selected study — https://github.com/st084331/Automatic_Steatosis_Recognition

FUNDING

This work was supported by St. Petersburg State University (project ID 94034084).

References

[1] M. Cohen, "The problem of liver steatosis and steatohepatitis in modern world," 2021.

[2] M. Naeem, M. R. P. Markus, M. Mousa, S. Schipf, M. Dörr, A. Steveling, A. Aghdassi, J.-P. Kühn, M.-L. Kromrey, M. Nauck, G. Targher, H. Völzke, and T. Ittermann, "Associations of liver volume and other markers of hepatic steatosis with all-cause mortality in the general population," Liver International, vol. 42, no. 3, pp. 575–584, 2022. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/liv.15133

[3] N. K. Desai, S. Harney, R. Raza, A. Al-Ibraheemi, N. Shillingford, P. D. Mitchell, and M. M. Jonas, "Comparison of controlled attenuation parameter and liver biopsy to assess hepatic steatosis in pediatric patients," The Journal of Pediatrics, vol. 173, pp. 160–164.e1, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S002234761600336X

[4] A. M. Pirmoazen, A. Khurana, A. E. Kaffas, and A. Kamaya, "Quantitative ultrasound approaches for diagnosis and monitoring hepatic steatosis in nonalcoholic fatty liver disease," Theranostics, vol. 10, pp. 4277 – 4289, 2020.

[5] J. Gu, S. Liu, S. Du, Q. Zhang, J. Xiao, Q. Dong, and Y. Xin, "Diagnostic value of MRI-PDFF for hepatic steatosis in patients with non-alcoholic fatty liver disease: a meta-analysis," Eur. Radiol., vol. 29, no. 7, pp. 3564–3573, Jul. 2019.

[6] J. Starekova, D. Hernando, P. J. Pickhardt, and S. B. Reeder, "Quantification of liver fat content with ct and mri: State of the art," Radiology, vol. 301, no. 2, pp. 250–262, 2021, pMID: 34546125.

[7] T. Poynard, V. Ratziu, S. Naveau, D. Thabut, F. Charlotte, D. Messous, D. Capron, A. Abella, J. Massard, Y. Ngo, M. Munteanu, A. Mercadier, M. Manns, and J. Albrecht, "The diagnostic value of biomarkers (SteatoTest) for the prediction of liver steatosis," Comparative Hepatology, vol. 4, no. 1, Dec. 2005. [Online]. Available: https://doi.org/10.1186/1476-5926-4-10

[8] R. Shrestha, S. KC, P. Thapa, A. Pokharel, N. Karki, and B. Jaishi, "Estimation of liver fat by FibroScan in patients with nonalcoholic fatty liver disease," Cureus, Jul. 2021. [Online]. Available: https://doi.org/10.7759/cureus.16414

[9] A. P. Gonchar, V. A. Gombolevskij, A. B. Elizarov, N. S. Kulberg, V. G. Klyashtorny, V. Y. Chernina, V. Y. Bosin, and S. P. Morozov, "Liver density in routine and low-dose computed tomography: the effect of image noise on measurement accuracy," Med. Vis., vol. 24, no. 1, pp. 39–47, May 2020.

[10] N. S. Kulberg, A. B. Elizarov, and V. S. Kovbas, "Program for liver image segmentation and liver x-ray density determination ctliverexam. certificate of state registration of the computer program no. 2019660983." 2019. [Online]. Available: https://www1.fips.ru/registers-doc-view/fips_servlet?DB=EVM&DocNumber=2019660983&TypeFile=html

[11] Z. Zhang, J. Weiss, J. Taron, R. Zeleznik, M. T. Lu, and H. J. W. L. Aerts, "Deep learning-based assessment of hepatic steatosis on chest ct," 2022. [Online]. Available: https://arxiv.org/abs/2202.02377

[12] P. M. Graffy, V. Sandfort, R. M. Summers, and P. J. Pickhardt, "Automated liver fat quantification at nonenhanced abdominal ct for population-based steatosis assessment," Radiology, vol. 293, no. 2, pp. 334–342, 2019, pMID: 31526254. [Online]. Available: https://doi.org/10.1148/radiol.2019190512

[13] H. S. Pettersen, I. Belevich, E. S. Røyset, E. Smistad, E. Jokitalo, I. Reinertsen, I. Bakke, and A. Pedersen, "Code-free development and deployment of deep segmentation models for digital pathology," 2021. [Online]. Available: https://arxiv.org/abs/2111.08430

[14] W. Li, F. Jia, and Q. Hu, "Automatic segmentation of liver tumor in CT images with deep convolutional neural networks," Journal of Computer and Communications, vol. 03, no. 11, pp. 146–151, 2015. [Online]. Available: https://doi.org/10.4236/jcc.2015.311023

[15] X. Chen, X. Wei, M. Tang, A. Liu, C. Lai, Y. Zhu, and W. He, "Liver segmentation in ct imaging with enhanced mask region-based convolutional neural networks," Annals of Translational Medicine, vol. 9, no. 24, 2021. [Online]. Available: https://atm.amegroups.com/article/view/85894

[16] "Visiopharm," https://visiopharm.com.

[17] "Halo ai," https://indicalab.com/halo-ai/.

[18] "Aiforia," https://www.aiforia.com.

[19] "Monai-lable," https://monai.io.

[20] "H-ai-l," https://github.com/SarderLab/H-AI-L.

[21] "Quickannotator," https://github.com/choosehappy/QuickAnnotator.

[22] "livermask," https://github.com/andreped/livermask.

[23] "scikit-learn," https://scikit-learn.org/stable/.

[24] "nibable," https://nipy.org/nibabel/.

[25] "scipy," http://www.scipy.org.

[26] "numpy," https://numpy.org.

[27] "Unifesp chest ct fatty liver competition," https://www.kaggle.com/competitions/unifesp-fatty-liver/overview.

[28] Q. H. Nguyen, H.-B. Ly, L. S. Ho, N. Al-Ansari, H. V. Le, V. Q. Tran, I. Prakash, and B. T. Pham, "Influence of data splitting on performance of machine learning models in prediction of shear strength of soil," Mathematical Problems in Engineering, vol. 2021, pp. 1–15, Feb. 2021. [Online]. Available: https://doi.org/10.1155/2021/4832864

[29] V. R. Joseph, "Optimal ratio for data splitting," Statistical Analysis and Data Mining: The ASA Data Science Journal, vol. 15, no. 4, pp. 531–538, Apr. 2022. [Online]. Available: https://doi.org/10.1002/sam.11583

[30] H. Abdi, "Multiple correlation coefficient," 2006.

[31] A. Schneider, G. Hommel, and M. Blettner, "Linear regression analysis," Deutsches Ärzteblatt international, Nov. 2010. [Online]. Available: https://doi.org/10.3238/arztebl.2010.0776

[32] Y. Kim and H. Oh, "Comparison between multiple regression analysis, polynomial regression analysis, and an artificial neural network for tensile strength prediction of BFRP and GFRP," Materials, vol. 14, no. 17, p. 4861, Aug. 2021. [Online]. Available: https://doi.org/10.3390/ma14174861

[33] Y. M. M. Eliseeva I. I., General Theory of Statistics: Textbook. Moscow: Finance and Statistics, 2002, vol. 4.