12th International Conference on Computer and Knowledge Engineering (ICCKE 2022), November 17-18, 2022, Ferdowsi University of Mashhad, Iran

Fatty Liver Level Recognition Using Particle Swarm Optimization (PSO) Image Segmentation and Analysis

Seyed Muhammad Hossein Mousavi
Independent Researcher
-
Tehran, Iran
mosavi.a.i.buali@gmail.com

Lyashenko Vyacheslav Victorovich
Kharkiv National University of Radio Electronics Media Systems and Technologies Department
Kharkiv, Ukraine
viacheslav.liashenko@nure.ua

Atiye Ilanloo
Faculty of Humanities- Psychology, Islamic Azad University of Rasht
Gilan, Iran
elanlooatiye@gmail.com

S. Younes Mirinezhad
Independent Researcher
-
Tehran, Iran
younes.mirinezhad@gmail.com

2022 12th International Conference on Computer and Knowledge Engineering (ICCKE) | 978-1-6654-7613-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCKE57176.2022.9960108

***Abstract*—Fatty liver or liver hepatic glycogen is one of the most common disorders of the liver nowadays. Clinical detection of this disorder by a human expert is increasing as our lifestyle leads us toward this phenomenon. So, making a fast and accurate expert system for fatty liver detection is essential in each clinic and that's why we intended to make one. Proposed expert system, works based on variety of image processing techniques and algorithms to detect fatty liver and recognize its level by four markers. Four segmentation techniques of Otsu, Watershed, K-Means and Particle Swarm Optimization (PSO) are employed to determine disorder level. Performance metrics of Accuracy, F-Score and IoU or Jaccard evaluated the precision of the proposed system. Finally, fatty liver level is calculated based on amount of fat deposits inside segmented image. Experiments are conducted on multiple data sample in high resolution with microscope zoom bigger or equal of 200 which are collected from the internet. All performance metrics and comparisons returned satisfactory results in comparing with traditional methods. Proposed system could achieve average accuracy value of 0.922 for all samples comparing with ground truth data. Additionally, F-Score and IoU performance metrics returned values are 0.872 and 0.907, respectively.**

Keywords— Fatty Liver Detection; Expert System; PSO; Image Segmentation; Fat Deposit; Hepatic Glycogen

I. Introduction

Accumulation of extra fat from food or alcohol in liver could cause fatty liver disease (hepatic steatosis). If it causes by heavy use of alcohol use then it calls Alcoholic Fatty Liver Disease (AFLD) and if causes by food then it calls Non-Alcoholic Fatty Liver Disease (NAFLD) [1]. This disease is wide spread nowadays, due to overuse of food and alcohol and less activity among people. So, extra fat doesn't burn enough and adds up to liver. Fatty liver normally doesn't show any symptom in early stages and if it continues without treatment, it may cause cirrhosis or liver cancer [2]. As it has fatal effect, research to find a proper solution is essential. Fatty liver shows signs of tiredness and pain in upper right side of the chest, which then it is time to visit the doctor. The doctor first asks the patient about symptoms and then prescribes, and in severe situation ask to liver function test blood. The problem is sometimes blood test fails to detect NAFLD and ultrasound scan is needed for robust detection.

That's when our proposed expert system comes forward to first detect fatty liver and then recognize its level by four level of markers using microscopic image data. It is possible to use Artificial Intelligence [3] techniques such as machine learning [4], pattern recognition [5], image processing [6] and evolutionary computation [7, 8] to make an expert system [9] for fatty liver detection. Main privileges of expert systems are their lack of error and tireless approach comparing with human expert. Multiple image processing techniques is employed in our method including nature-inspired segmentation [10] for achieving more precise results.

Paper is consisted of five main parts starting with Introduction which explains basics. Second part is all about prior related researches in the subject area. Section 3 pays to proposed method in details. Section 4 investigates all evaluations and results in the experiment and finally, conclusion, future works and suggestion makes the fifth section.

II. Literature Review

One of the mentionable types of research in fatty liver detection using image segmentation is Guo, Xiaoyuan, et al, research in 2019 [11]. They used Mask-RCNN algorithm to segment fat deposits in liver samples and achieved 78.87 % average accuracy with their method.

978-1-6654-7613-3/22/$31.00 ©2022 IEEE

Another research in these areas belongs to Zhang, Qin-He, et al in 2021 [12]. They used Region of Interest (ROI) on Magnetic Resonance Imaging (MRI) in order to segment liver images for NAFLD and received promising results.

Kullberg, Joel, et al [13], in 2017 used Computed Tomography (CT) images belongs to 107 subjects in 3 slices to estimate fat deposits in patients by common segmentation techniques and achieved high correlation of 0.930 for their analysis.

Rhyou, Se-Yeol, and Jae-Chern Yoo in 2021 [14], presents a fully automatic liver steatosis prediction model using three deep neural networks and semantic segmentation to achieved 99.78 % accuracy.

Nativ, Nir I., et al made a system for fatty liver detection using segmentation techniques in 2014 [15]. They used 54 images from 9 subjects and could achieve 99.3 % accuracy in fatty liver detection by their system.

Significant research to characterize normal, fatty and heterogeneous liver, using textural analysis of liver Ultrasound images introduced by Owjimehr, Mehri, et al, in 2015 [16]. Their system selects region of interest by segmentation and uses wavelet transform to extract statistical features and discriminates the type of diseases in liver including fatty liver. They could get 100 % accuracy on their data. The need for strong single image-based segmentation is vivid which need to be satisfied.

III. Proposed Method

There are four markers which based on those, the level of fatty liver is determined in this research. These markers are presented in Figure 1 and defined by a human expert. This research uses, high resolution microscopic images with zoom higher than 200. Figure 2. Represents some samples which are used in the experiment. Figure 3 shows liver surface in different stages. Also, all data are collected from the internet. The novelty of this research lays behind using a nature-inspired algorithm for segmentation on single images which outperforms traditional single image-based methods.

- Workflow

At the beginning, liver image investigates and analysis by human expert and send to computer system for automatic recognition. Some pre-processing steps such as intensity adjustment [17], histogram equalization [17] and canny edge detection [18] apply on input image for improvements and better recognition process. There are four segmentation [19] techniques in this research four marker selection namely, Otsu [20], Watershed [21], K-Means [22] and Particle Swarm Optimization (PSO) segmentations [23]. We didn't compare proposed method with deep learning approaches as our method is single mage-based and deep learning is multiple image-based. In order to evaluate the performance of the segmentations, three performance metrics namely, Accuracy [24], F-Score [25], and IoU (Jaccard) [26] are employed. Final step is consisted of fatty liver level recognition based on fat deposits amount and marker selection process. Clearly, marker and its level are based on Figure 1 content. Figure 4 depicts the flow char of the proposed method. In tensity adjustment is to map lowest and highest intensity of image into minimum and maximum intensity of the full intensity range. For instance, a min and max intensity of 0 and 255 for an 8-bit image would have contrast adjustment or intensity adjustment of (1).

$$f_{\mathrm{ac}}(a) = (a - a_{\mathrm{low}}) \cdot \frac{255}{a_{\mathrm{high}} - a_{\mathrm{low}}} \quad (1)$$

Let f be a given image represented as a m_r by m_c matrix of integer pixel intensities ranging from 0 to $L-1$. L is the number of possible intensity values. Let p denote the normalized histogram of f with a bin for each possible intensity. So (2) is:

$$p_n = \frac{\text{number of pixels with intensity } n}{\text{total number of pixels}} \; n = 0,1,\ldots,L-1. \quad (2)$$

The histogram equalized image g will be defined by (3):

$$g_{i,j} = \mathrm{floor}\left((L-1)\sum_{n=0}^{f_{i,j}} p_n\right) \quad (3)$$

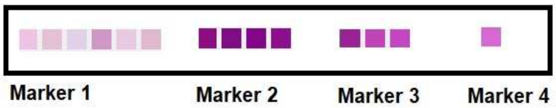

Fig. 1. Markers for fatty liver level detection

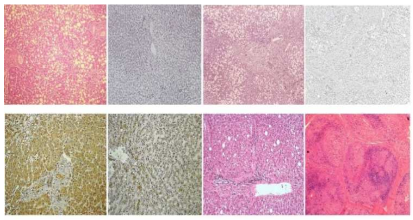

Fig. 2. Some samples employed inside the experiment

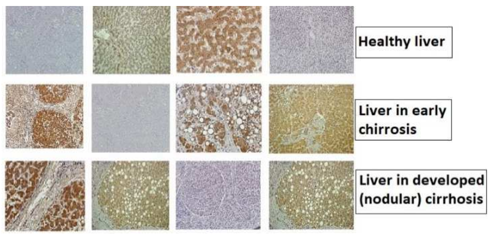

Fig. 3. Liver surface in different health stages

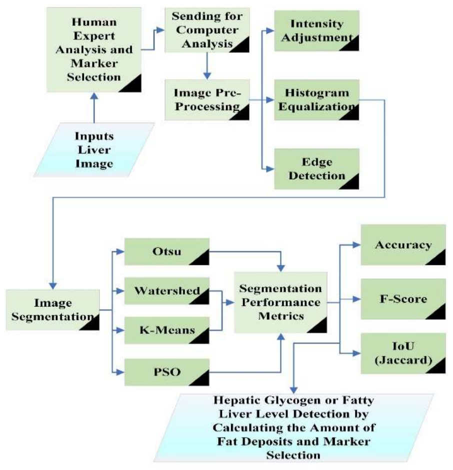

Fig. 4. Flowchart of the proposed method

Clearly K-means [27] algorithm could cluster close groups of samples into different number of clusters based on their distance in the space. Same method could be employed for image pixel values which, could be gray level, intensity or color. One of the best implementations for K-means on images belongs to Dhanachandra, N, et al in 2015 [28]. Obviously, this method is categorized into cluster-based segmentations methods.

Another example is Watershed image segmentation algorithm [21]. It is a region-based method and this algorithm is based on drainage in basins or rivers. It works based on the water level as water height represents the separating line between one region or segment to another. A commonly used segmentation technique is called Otsu's thresholding segmentation [20] which is categorized in threshold-based segmentations. It works as a thresholding process for pixel intensity values in order to minimizing interclass variance for separating pixels into two classes of foreground and background. Number of threshold value, determines, number of segments. PSO image segmentation pseudo code is as follow.

Figure 5 presents an example of fatty liver level recognition. Figure 6 presents important steps of proposed method on a real sample. In PSO segmentation, process improves during iterations based on number of segments, number of initial population and iterations. Parameters of the experiment are as follow: iterations of 250, population 50, 4 segments, inertia coefficient of 1, damping ration of 0.99, and personal learning coefficient of 2. Figure 7 shows different markers on example. Figure 8 illustrates another example of PSO segmentation with boundaries. Figure 9 depicts, an experiment sample with alongside with the ground truth image for comparison. Additionally, Figure 10 and 11 shows two other results with proposed segmentation system.

PSO Segmentation Pseudo Code

Start
Training Using Particle Swarm Optimization (Input: Pre-Processed Image)
Goal: To Segment the Input Image
Initialize the Population Size N and Number of Generations
While *(number of generations is not reached)*
Initialized Particles with Random Position and Velocity for PSO
Evaluate the Fitness of Particles for each Pixel and Their Corresponding Distance for PSO
Find and update pbest and gbest for PSO
Calculate and Update Velocity and Position for PSO
Show gbest the Optimal Solution for PSO
Update the Best Solution Found for Pixel and Distance by PSO
End While
Apply Best Clusters Found on Image to Segment
Output: Evolutionary Segmented Image
End

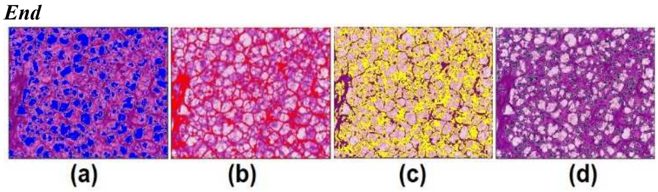

Fig. 5. (a) marker 1 (blue), (b) marker 2 (red), (c) marker 3 (yellow), (d) marker 4 (some black spots), the result is marker 2 with 21.5%

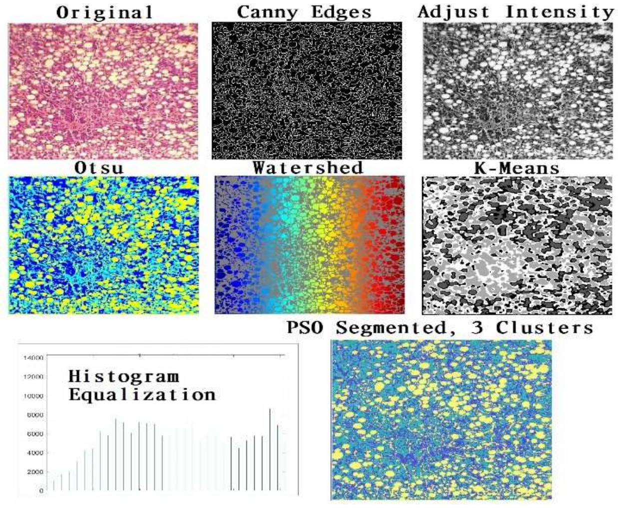

Fig. 6. Proposed method on a real sample

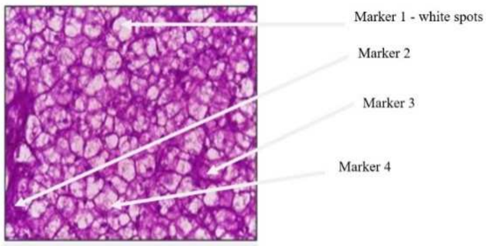

Fig. 7. Markers on an example

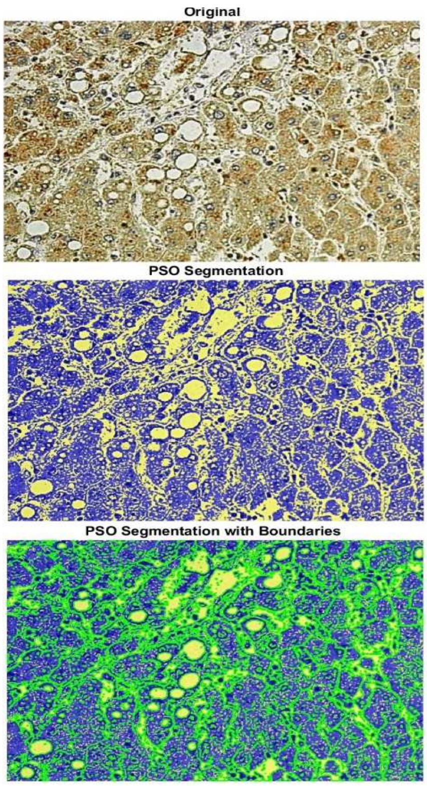

Fig. 8. Another experiment result with PSO and boundaries around fat deposits

In Figure 7, The area for the marker 1 is 22.6%, The area for the marker 2 is 21.5%, The area for the marker 3 is 43,2%, and the area for the marker 4 – 12.7%. If marker 2, marker 3 and marker 4 are hepatic glycogen, the area of hepatic glycogen is 77.4%. In view of the first and second approaches hepatic glycogen area is within the: 77.4% – 81.96%. (The result depends on the exact choice of marker). Marker 2 - is fully hepatic glycogen, Marker 3 - hepatic glycogen is 80% and Marker 4 - hepatic glycogen is 50%. Then the area of hepatic glycogen is:

$$21.5 * 1 + 43.8 * 0.8 + 12.7 * 0.5 = 21.5 + 35.04 + 6.35 = 62.89\ \% \tag{4}$$

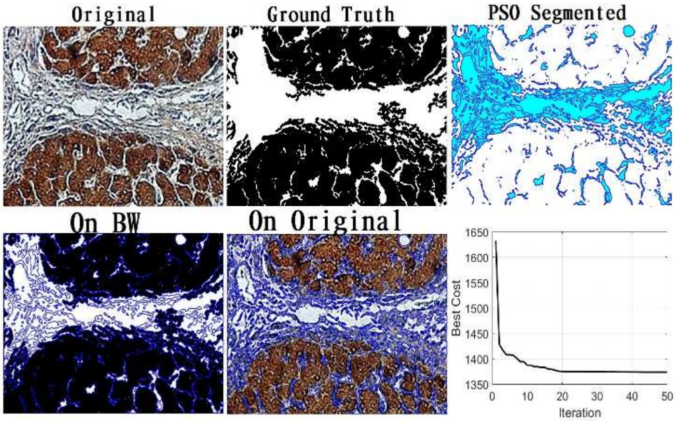

Fig. 9. An Experiment sample alongside with the ground truth and the final segmentation result

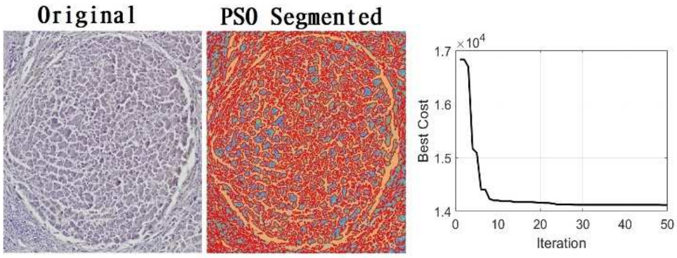

Fig. 10. Original image and PSO segmented result on a selected experiment sample (1)

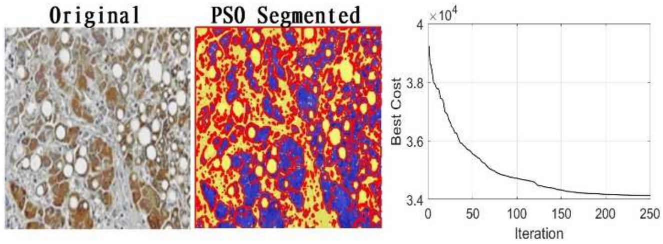

Fig. 11. Original image and PSO segmented result on a selected experiment sample (2)

IV. Evaluations and Results

In order to evaluate the performance of the proposed segmentation method, three most famous performance metrics of accuracy [29], boundary F1 score [30] and Intersection over Union (IoU) [31] metrics are employed which, has application in comparison step too. Accuracy simply is the ratio of correctly classified pixels for objects.

$$Accuracy\ Score\ =\ Number\ of\ True\ Positive\ /\ (Number\ of\ True\ Positive\ +\ Number\ of\ False\ Negative) \tag{5}$$

The boundary F1 (BF) contour matching score indicates how well the predicted boundary of each class aligns with the true boundary. The BF score is defined as the harmonic mean (F1-measure) of the precision and recall values with a distance error tolerance.

$$F\ score\ =\ 2 * precision * recall\ /\ (recall\ +\ precision) \tag{6}$$

Intersection over union (IoU), also known as the Jaccard similarity coefficient, is the most commonly used metric for segmentation. IoU is a statistical accuracy measurement that penalizes false positives. For each class, IoU is the ratio of correctly classified pixels to the total number of ground truth and predicted pixels in that class.

$$IoU\ score\ =\ True\ Positive\ /\ (True\ Positive\ +\ False\ Positive\ +\ False\ Negative) \tag{7}$$

Tables I to III presents performance metrics results on all image samples comparing with ground truth data. Figure 12 shows Tables I to III as box plot. Figures 13 and 14 depict some examples of proposed level recognition system alongside with related distribution and final recognition percentage.

TABLE I. ACCURACY PERFORMANCE METRICS RESULTS ON ALL SAMPLE IMAGES AND COMPARISONS

Accuracy	**Otsu**	**Watershed**	**K-Means**	**PSO**
Sample Image 1 – Healthy	0.8398	0.6558	0.7562	**0.9465**
Sample Image 2 – Healthy	0.8106	0.7961	0.8657	**0.9070**
Sample Image 3 – Healthy	0.8062	0.6262	0.8456	**0.9659**
Sample Image 4 – Healthy	0.8546	0.6136	0.8729	**0.9281**
Sample Image 5 - Early Chirrosis	0.7948	0.6603	0.6392	**0.9011**
Sample Image 6 - Early Chirrosis	**0.8816**	0.5832	0.7107	0.8107
Sample Image 7 - Early Chirrosis	0.8283	0.6265	0.6817	**0.8604**
Sample Image 8 - Early Chirrosis	0.8689	0.5842	0.7956	**0.9547**
Sample Image 9 - Developed Cirrhosis	0.9310	0.5216	0.7036	**0.9875**
Sample Image 10 - Developed Cirrhosis	0.8906	0.5610	0.6735	**0.9159**
Sample Image 11 - Developed Cirrhosis	0.8683	0.6130	0.7221	**0.9174**
Sample Image 12 - Developed Cirrhosis	0.8751	0.6055	0.7966	**0.9761**

TABLE II. F-SCORE PERFORMANCE METRICS RESULTS ON ALL SAMPLE IMAGES AND COMPARISONS

F-Score	Otsu	Watershed	K-Means	PSO
Sample Image 1 – Healthy	0.7144	0.7956	0.7882	**0.8253**
Sample Image 2 – Healthy	0.7193	0.6178	0.8056	**0.9399**
Sample Image 3 – Healthy	0.7746	0.5705	0.7818	**0.9099**
Sample Image 4 – Healthy	0.7684	0.6457	0.6256	**0.8142**
Sample Image 5 - Early Chirrosis	**0.8236**	0.5282	0.6962	0.7254
Sample Image 6 - Early Chirrosis	0.7661	0.5914	0.6304	**0.8202**
Sample Image 7 - Early Chirrosis	0.8613	0.5602	0.7833	**0.9363**
Sample Image 8 - Early Chirrosis	**0.9073**	0.4799	0.6986	0.8972
Sample Image 9 - Developed Cirrhosis	0.7961	0.5119	0.6246	**0.8821**
Sample Image 10 - Developed Cirrhosis	0.7783	0.5761	0.7112	**0.8394**
Sample Image 11 - Developed Cirrhosis	0.8361	0.5813	0.7562	**0.9665**
Sample Image 12 - Developed Cirrhosis	0.7251	0.8110	0.8001	**0.9108**

TABLE III. IOU PERFORMANCE METRICS RESULTS ON ALL SAMPLE IMAGES AND COMPARISONS

IoU	Otsu	Watershed	K-Means	PSO
Sample Image 1 – Healthy	0.7276	**0.8898**	0.8838	0.8828
Sample Image 2 – Healthy	0.7253	0.6413	0.8409	**0.9022**
Sample Image 3 – Healthy	0.7761	0.5748	0.7987	**0.9748**
Sample Image 4 – Healthy	0.8416	0.7105	0.6707	**0.8689**
Sample Image 5 - Early Chirrosis	**0.8532**	0.6027	0.7151	0.7941
Sample Image 6 - Early Chirrosis	0.7845	0.6282	0.6939	**0.8982**
Sample Image 7 - Early Chirrosis	0.8694	0.6531	0.8609	**0.9859**
Sample Image 8 - Early Chirrosis	**0.9509**	0.5246	0.7292	0.9481
Sample Image 9 - Developed Cirrhosis	0.8472	0.5937	0.7041	**0.9465**
Sample Image 10 - Developed Cirrhosis	0.8162	0.6573	0.7645	**0.8745**
Sample Image 11 - Developed Cirrhosis	0.9301	0.6689	0.8112	**0.9387**
Sample Image 12 - Developed Cirrhosis	0.7731	0.8164	0.8183	**0.8724**

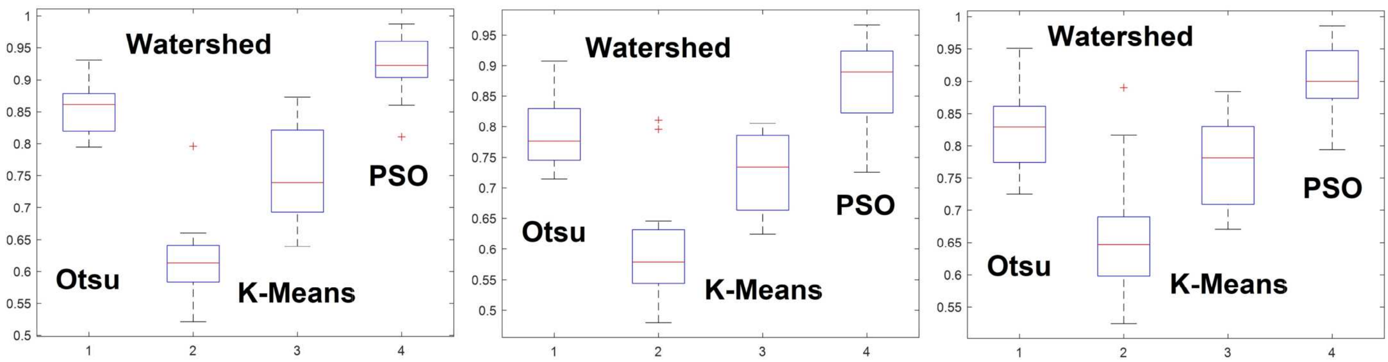

Fig. 12. Results from Tables 1 to 3 as box plot (left: Accuracy, center: F-Score and right: IoU)

Looking at Tables I to III, PSO segmentation method has superiority over other three and Otsu method took second place. Third place belongs to K-means segmentation and Watershed placed at end of the ranking.

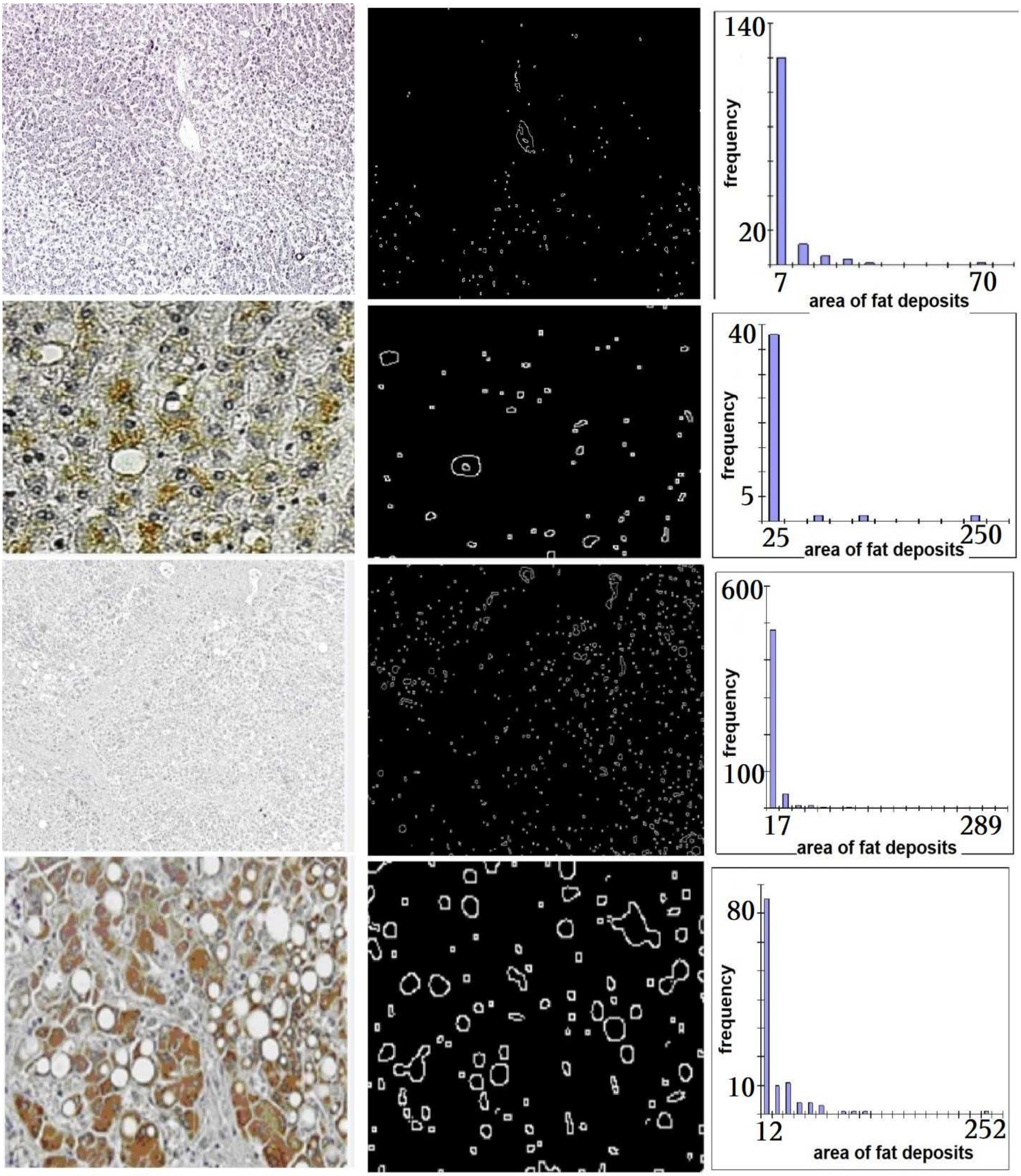

Fig. 13. First row: the proportion of fat deposits is 0.18%, second row: the proportion of fat deposits is 1.18%, third row: the proportion of fat deposits is 2.67%, fourth row: the proportion of fat deposits is 6.81% plus Distribution fat deposits (last column)

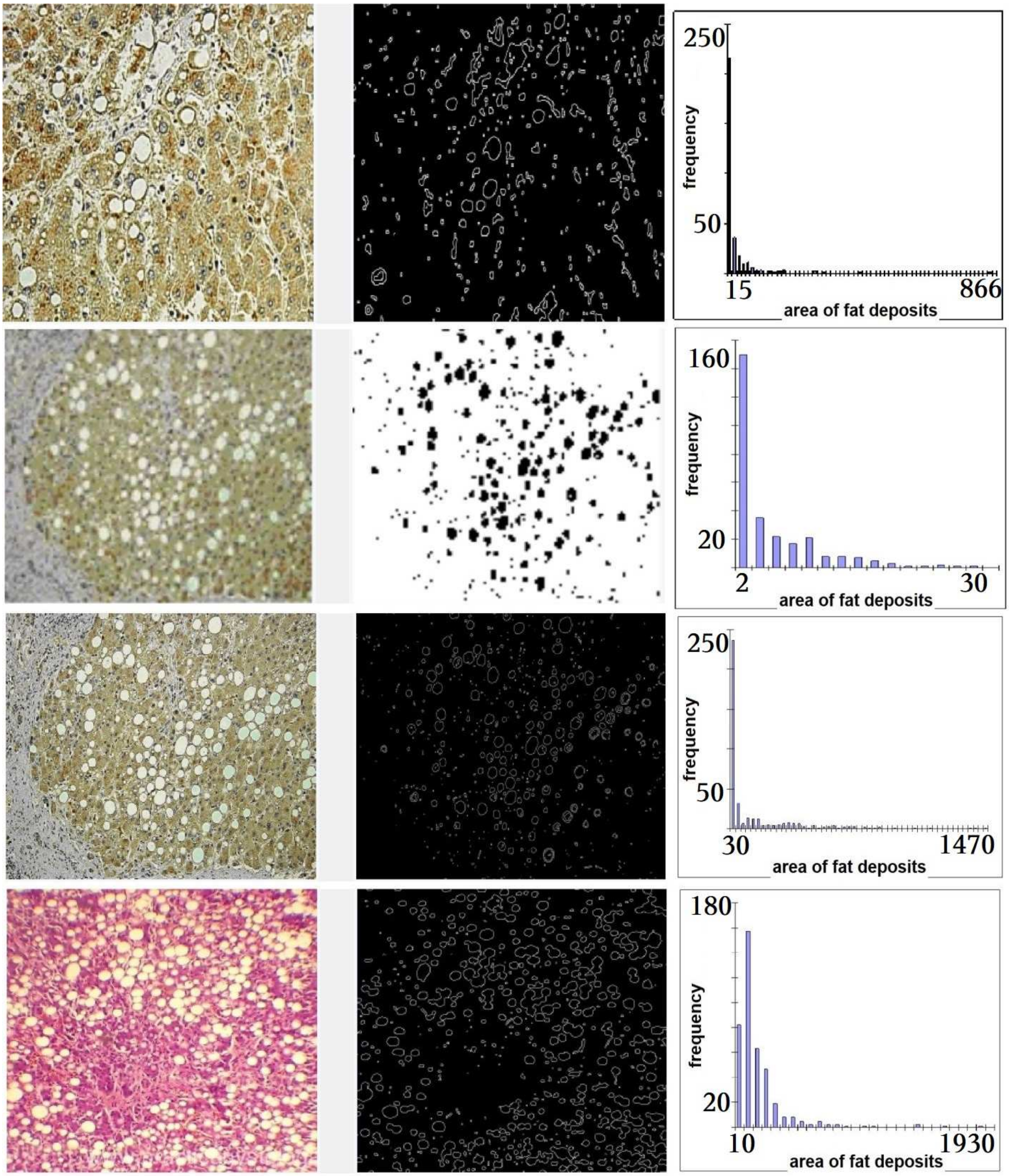

Fig. 14. First row: The proportion of fat deposits is 4.92%, second row: The proportion of fat deposits is 10,2%, third row: The proportion of fat deposits is 11.3%, fourth row: The proportion of fat deposits is 21.75% plus Distribution fat deposits (last column)

V. Conclusion, Future Works and Suggestions

By employing nature-inspired algorithms in image processing and especially image segmentation, it is possible to achieve even more precise result from segmentation as this research proved PSO algorithm could reach to that point. Clearly, this research could increase fatty liver level recognition accuracy and speed by sampling from patient's liver which could be replaced by human expert. Required results from performance metrics showed better performance in Otsu and PSO segmentation techniques. Even by changing hyperparameters of traditional methods, PSO segmentations returned superiority over all of them. It can be concluded that, proposed method could be used in both AFLD and NAFLD diseases. Using other nature-inspired algorithm such as Genetic Algorithm (GA) and others in segmentation techniques is of future works. It is suggested to employ proposed method on different data with higher or lower microscopic zoom.

References

[1] Clark, Jeanne M., Frederick L. Brancati, and Anna Mae Diehl. "Nonalcoholic fatty liver disease." Gastroenterology 122.6 (2002): 1649-1657.

[2] Nault, Jean-Charles, et al. "The role of telomeres and telomerase in cirrhosis and liver cancer." Nature Reviews Gastroenterology & Hepatology 16.9 (2019): 544-558.

[3] Winston, Patrick Henry. Artificial intelligence. Addison-Wesley Longman Publishing Co., Inc., 1992.

[4] Dezfoulian, Mir Hossein, et al. "Optimization of the Ho-Kashyap classification algorithm using appropriate learning samples." 2016 Eighth International Conference on Information and Knowledge Technology (IKT). IEEE, 2016.

[5] Al-Sharo, Y. M. "Neural Networks As A Tool For Pattern Recognition of Fasteners." International Journal of Engineering Trends and Technology 69.10 (2021): 151-160.

[6] Akhmetov, V., et al.. "Astrometric reduction of the wide-field images." Advances in Intelligent Systems and Computing IV: Springer Nature Switzerland. 1080 (2020): 896–909. DOI: 10.1007/978-3-030-33695-0_58.

[7] Mousavi, Seyed Muhammad Hossein, Vincent Charles, and Tatiana Gherman. "An evolutionary Pentagon Support Vector finder method." Expert Systems with Applications 150 (2020): 113284.

[8] Mousavi, Seyed Muhammad Hossein, S. Younes MiriNezhad, and Mir Hossein Dezfoulian. "Galaxy gravity optimization (GGO) an algorithm for optimization, inspired by comets life cycle." 2017 Artificial Intelligence and Signal Processing Conference (AISP). IEEE, 2017.

[9] Mousavi, Seyed Muhammad Hossein, et al. "A PSO fuzzy-expert system: As an assistant for specifying the acceptance by NOET measures, at PH. D level." 2017 Artificial Intelligence and Signal Processing Conference (AISP). IEEE, 2017.

[10] Larabi-Marie-Sainte, Souad, Reham Alskireen, and Sawsan Alhalawani. "Emerging applications of bio-inspired algorithms in image segmentation." Electronics 10.24 (2021): 3116.

[11] Guo, Xiaoyuan, et al. "Liver steatosis segmentation with deep learning methods." 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE, 2019.

[12] Zhang, Qin-He, et al. "Hepatic fat quantification of magnetic resonance imaging whole-liver segmentation for assessing the severity of nonalcoholic fatty liver disease: comparison with a region of interest sampling method." Quantitative Imaging in Medicine and Surgery 11.7 (2021): 2933.

[13] Kullberg, Joel, et al. "Automated analysis of liver fat, muscle and adipose tissue distribution from CT suitable for large-scale studies." Scientific reports 7.1 (2017): 1-11.

[14] Rhyou, Se-Yeol, and Jae-Chern Yoo. "Cascaded Deep Learning Neural Network for Automated Liver Steatosis Diagnosis Using Ultrasound Images." Sensors 21.16 (2021): 5304.

[15] Nativ, Nir I., et al. "Automated image analysis method for detecting and quantifying macrovesicular steatosis in hematoxylin and eosin–stained histology images of human livers." Liver Transplantation 20.2 (2014): 228-236.

[16] Owjimehr, Mehri, Habibollah Danyali, and Mohammad Sadegh Helfroush. "An improved method for liver diseases detection by ultrasound image analysis." Journal of Medical Signals and Sensors 5.1 (2015): 21.

[17] Wang, Yuwei, et al. "Nonlinear Correction for Fringe Projection Profilometry with Shifted-Phase Histogram Equalization." IEEE Transactions on Instrumentation and Measurement 71 (2022): 1-9.

[18] Mousavi, Seyed Muhammad Hossein, Vyacheslav Lyashenko, and Surya Prasath. "Analysis of a robust edge detection system in different color spaces using color and depth images." Компьютерная оптика 43.4 (2019): 632-646.

[19] Rabotiahov, Andrii, et al. "Bionic image segmentation of cytology samples method." 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET). IEEE, 2018.

[20] Zhang, Jun, and Jinglu Hu. "Image segmentation based on 2D Otsu method with histogram analysis." 2008 international conference on computer science and software engineering. Vol. 6. IEEE, 2008.

[21] Ju, Aiyun, and Zhongli Wang. "A novel fully convolutional network based on marker-controlled watershed segmentation algorithm for industrial soot robot target segmentation." Evolutionary Intelligence (2022): 1-18.

[22] Zhang, Haixia, and Qingxiu Peng. "PSO and K-means-based semantic segmentation toward agricultural products." Future Generation Computer Systems 126 (2022): 82-87.

[23] Zhang, Yong, et al. "Image segmentation using PSO and PCM with Mahalanobis distance." Expert systems with applications 38.7 (2011): 9036-9040.

[24] Popovic, Aleksandra, et al. "Statistical validation metric for accuracy assessment in medical image segmentation." International Journal of Computer Assisted Radiology and Surgery 2.3 (2007): 169-181.

[25] Sokolova, Marina, Nathalie Japkowicz, and Stan Szpakowicz. "Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation." Australasian joint conference on artificial intelligence. Springer, Berlin, Heidelberg, 2006.

[26] Setiawan, Agung W. "Image Segmentation Metrics in Skin Lesion: Accuracy, Sensitivity, Specificity, Dice Coefficient, Jaccard Index, and Matthews Correlation Coefficient." 2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM). IEEE, 2020.

[27] Mousavi, Seyed Muhammad Hossein, S. Younes MiriNezhad, and Atiye Mirmoini. "A new support vector finder method, based on triangular calculations and K-means clustering." 2017 9th International Conference on Information and Knowledge Technology (IKT). IEEE, 2017.

[28] Dhanachandra, Nameirakpam, Khumanthem Manglem, and Yambem Jina Chanu. "Image segmentation using K-means clustering algorithm and subtractive clustering algorithm." Procedia Computer Science 54 (2015): 764-771.

[29] Popovic, Aleksandra, et al. "Statistical validation metric for accuracy assessment in medical image segmentation." International Journal of Computer Assisted Radiology and Surgery 2.3 (2007): 169-181.

[30] Sokolova, Marina, Nathalie Japkowicz, and Stan Szpakowicz. "Beyond accuracy, F-score and ROC: a family of discriminant measures for

performance evaluation." Australasian joint conference on artificial intelligence. Springer, Berlin, Heidelberg, 2006.

[31] Setiawan, Agung W. "Image Segmentation Metrics in Skin Lesion: Accuracy, Sensitivity, Specificity, Dice Coefficient, Jaccard Index, and Matthews Correlation Coefficient." 2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM). IEEE, 2020.