Journal of Medical Systems (2019) 43: 101
https://doi.org/10.1007/s10916-019-1207-7

IMAGE & SIGNAL PROCESSING

The Influence of Different Segmentation Methods on the Extraction of Imaging Histological Features of Hepatocellular Carcinoma CT

Sen Zhao[1] · Wenyan Ren[1] · Yan Zhuang[1] · Zhixue Wang[1]

Received: 26 December 2018 / Accepted: 11 February 2019 / Published online: 14 March 2019

Abstract
In order to analyze the influence of different segmentation techniques on hepatocellular carcinoma (HCC) CT (Computed Tomography) imaging histological feature extraction, Grow Cut method and Graph Cut method are used to segment hepatocellular carcinoma from arterial CT images of HCC patients, and the stability and repeatability of imaging histological features are studied. Meanwhile, hierarchical clustering method is used to reduce the redundancy of features. The results show that the repeatability and redundancy mainly depend on the method of tumor segmentation. Semi-automatic segmentation method can improve the repeatability of image features, and hierarchical clustering can reduce the redundancy of features. Different segmentation techniques have different effects on the extraction of histological features of CT images of HCC.

Keywords Segmentation · HCC · Imaging histological features · Repeatability

Introduction

According to the latest data released by international research institutes affiliated to the World Health Organization (WHO), the incidence of liver cancer is high in underdeveloped areas. Statistics show that at least 550,000 people die of hepatocellular carcinoma (HCC) every year [1, 2]. According to death reports, 57% of the cases are in less developed countries along the coast of Southeast Asia and the Pacific, and some parts of China are also among the high incidence areas of HCC in the world. HCC ranks second among fatal cancers worldwide, with a conservative estimate of an overall mortality rate of 0.95 [3]. Fortunately, the progress and development of medical science and technology have provided more and more effective ways for the treatment of HCC, so that the vast number of patients see hope. Therefore, it is particularly important to detect liver lesions as early as possible and to formulate reasonable and effective treatment plans.

Computed tomography (CT) can solve the problems of blurred X-ray imaging and poor tissue resolution. It can clearly see the structure of soft tissue, and solve the imaging problem of difficult X-ray parts. At the same time, with the progress of tube detector technology and the popularization of low-dose technology, it greatly improves the accuracy and safety of clinical diagnosis. Because of the high definition of CT images, obvious imaging of soft tissue can highlight the pathological organs, and the price is moderate, so that it has been widely used in the diagnosis of liver diseases.

The purpose of medical image segmentation is to extract interested regions from medical images, such as some organs, tissues or lesions. Computer-aided diagnosis, three-dimensional visualization, graphics-guided surgery, virtual endoscopy and so on are realized by image processing technology [4]. Liver is accurately and reliably segmented from abdominal CT images, which is the first step in the early diagnosis of liver diseases, estimation of liver size, and establishment of 3D model and it is also a key step [5]. Usually, in practical clinical applications, the liver is segmented manually from CT images by doctors with relevant practical experience and expertise [6]. However, this process is very time-consuming and energy-consuming. Moreover, influenced by doctors' subjective factors, different doctors often get different results.

This article is part of the Topical Collection on *Image & Signal Processing*

✉ Zhixue Wang
zhixuewang666@126.com

[1] Medical Imaging Center, The First Affiliated Hospital of Henan University, No.357 Ximen Street, Kaifeng City 475001, Henan, China

Image acquisition and tumor segmentation

CT image acquisition

Primary liver tumors are the most common malignant tumors in the liver [7]. Conventional single-phase CT scans commonly used in clinical diagnosis are mainly in the portal phase of the liver. The detection rate of primary liver cancer is low [8]. Hepatic artery blood supply is the main cause of primary liver cancer. Scanning in arterial phase can improve the detection rate of primary liver cancer. For hepatic hemangioma, portal vein is the main blood supply. Portal phase enhancement can significantly improve the detection rate.

Because the duration of hepatic artery phase is short, only 20–30 s, arterial phase enhanced scanning can show the characteristics of hepatocellular carcinoma more clearly. There are two ways of blood supply to the liver: hepatic artery (about 20–25%) and portal vein (about 75–80%). Contrast medium can flow into the liver through the above two pathways. According to the time-density curve of hepatic parenchyma and abdominal aorta, the enhancement scanning process of liver is divided into three stages: arterial phase (25–30 s after contrast medium injection), portal phase (60–70 s after contrast medium injection) and equilibrium phase (120 s after contrast medium injection) [9].

Arterial phase enhanced CT images have a strong advantage in detecting primary HCC [10], and can better reflect the blood supply characteristics of the tumor area than other enhanced CT images. Therefore, 15 cases of primary HCC patients with arterial phase enhanced CT images are obtained, which can better capture the texture features of primary HCC, and obtain more image histology features information.

Three-dimensional liver segmentation based on CT images

The three-dimensional liver segmentation method developed has three main steps: first, edge detection of differential images; second, using edge detection operator to get the initial contour of the liver and re-detect to determine the final contour of the liver; third, deleting the adhesion area and integrating the lesion area.

Here, differential processing is introduced before edge detection in order to reduce the impact of the lesion area on liver segmentation results. In CT images, the brightness of liver and lesion areas in plain scan phase is often lower than that in portal phase. In differential images, the brightness of liver and lesion areas is lower than 0, which is taken 0 here. However, the brightness of other areas outside the liver region in the differential image is usually higher than 0, so according to this characteristic, the liver region can be detected by using the differential image.

Three-dimensional label processing is used to segment the initial liver region from the edge-detected image. Then, the slices with a total thickness of 75 mm at 25 mm away from the heart are selected as the starting point, and these selected binary images are used to calculate the average liver. Each of the sequence images of the average liver and the initial liver region are used to calculate the average liver. If the difference is the smallest, the slice is considered to be the largest slice SO in the sequence image of the initial liver region. Since all parts of the original image are homogeneous in the liver region, but not in the region outside the liver, brightness can be selected as a threshold to further confirm the liver region. As mentioned above, the largest slice SO in the original image liver region can be found, and then the average density Gavr of the slice SO can be calculated. The liver region of the original image can be confirmed by using Gavr±SD as the threshold.

In liver segmentation, parts of the kidney and the heart are often segmented together. This is because the brightness of kidney and heart overlaps with that of liver. It is not possible to delete the adhesion parts only by portal phase. However, these adhesions have different gray values in arterial phase and liver area and can successively and successfully delete the adhesive parts. In the process of liver segmentation, the brightness of hemangioma and liver is totally different. Although hemangioma is preserved in edge detection, it is deleted when the liver region is further confirmed. In order to find out the area of hemangioma, the average liver region is adopted to differentiate from the initial three-dimensional liver region. The hemangioma regions are selected by using roundness and volume. Then, the average gray value of the selected hemangioma area is calculated and selected. The hemangioma area is further confirmed by using threshold technology. Finally, the hemangioma area is integrated with the previously segmented liver area and the complete liver area is obtained.

Tumor segmentation for primary HCC

Grow Cut is a competitive region growing algorithm based on cellular automata. It has good performance in speed and accuracy and it is also an effective interactive segmentation method. During the implementation of the algorithm, the segmentation problem is redefined as a clustering problem, and a fast and approximate solution is worked out. It has good effect in three-dimensional medical image segmentation application.

Grow Cut in 3D Slice is achieved through Fast Grow Cut Effect module and the implementation steps are shown as follows:

Step 1: load CT image;
Step 2: mark the tumour and non-tumour regions and give a series of initial marker points;
Step 3: run Grow Cut, obtain the initial segmentation results, and modify them if they are unsatisfactory.

Step 4: continue running Grow Cut until the partition is complete.

In 3D Slicer, the segmentation method of Graph Cut based on star-shape priori is provided. It is realized through Graph Cut Interactive Segmenter module and the implementation steps are shown as follows:

Step 1: load CT image;
Step 2: two benchmarks are used to locate the center of the tumour at the beginning and end of the tumor, respectively. The other two benchmarks are located at the diagonal point of the rectangle containing the tumor at the maximum level of the tumor in the middle.
Step 3: check 3D or 2D star-shape constraints as needed;
Step 4: Graph Cut automatically generates regions of interest according to the selection of benchmark points, and continues to run to get the initial segmentation results. If it is unsatisfactory, modify them.
Step 5: continue running Graph Cut until the partition is complete.

Imaging histological feature extraction

The core content of imaging histology is to quantitatively describe the phenotypic information of lesions by extracting the high-dimensional feature data in the region of interest (tumor). Therefore, when the tumor area is determined, the imaging histological features can be extracted. At present, some open source and commercial software can be used in imaging histology research, which greatly promotes the clinical application and development of imaging histology. Open source software includes IBEX (Imaging Biomarker Explorer), CGITA (Chang Guang Image Texture Analysis) and MaZda. Commercial software includes Radiomi X™ and TexRAD™. In addition, there are some image feature extraction tools based on MATLAB (matrix and laboratory), including software adapted from CERR (Computational Environment for Radiotherapy Research). The extracted imaging histological features include intensity-based first-order statistical features, geometry-based features and texture-based features. In addition, there are some filter-based imaging histological features.

The features of image histology include intensity-based histogram feature, texture feature and geometry feature.

Strength histogram feature, also known as first-order statistical feature, is the first-order histogram feature based on the intensity distribution of descriptive image pixels. For CT images, pixel strength refers to CT value (HU), MR image refers to signal strength, and PET image refers to standard uptake value. Gray intensity histogram shows the distribution of the gray level of the pixels in the image. According to the size of the pixel value, the frequency of its occurrence is counted, that is, the function of the distribution of the pixel gray level. Gray intensity histogram describes the one-dimensional information of an image. It only reflects the distribution of the gray level of a pixel, but does not care about its spatial position. The gray intensity histogram formula of gray image X is as follows:

$$H(i) = \frac{n(i)}{N}, i = 0, 1, \cdots, 2^k \quad (1)$$

In Formula (1), $H(i)$ refers to the frequency of occurrence of gray level i, n(i) indicates the number of pixels with the gray level of i, N suggests the total number of pixels in the image, and k is the bit depth of image.

Gray histogram features provide the distribution information of the intensity of the pixels in the image, and do not involve the spatial position relationship between the pixels. In order to quantify the position relationship and texture distribution among pixels, three techniques are often introduced: Gray-level co-occurrence matrix (GLOM), gray-level run length matrix (GLRLM) and neighbourhood gray-tone difference matrix (NGTDM). Before calculating texture features, in order to avoid the influence of noise and sparse matrix, the original images are resampled into 8-bit gray images.

GLOM is a statistical method established based on estimating the probability density function of the second-order combination condition of the image. It can reflect the comprehensive information of the gray level of the image about direction, adjacent interval and change range. It is the basis of analyzing the local pattern of the image and their arrangement rules. It is defined as: within the image domain, the probability of two pixels with distance d and direction e appearing in the image.

Shape features are quantitative descriptions of the geometry of tumors, which can usually be extracted from the two-dimensional and three-dimensional information of tumors. There are mainly two kinds of representation methods for shape features, contour features and region features. The first type mainly focuses on the boundaries of regions of interest, while the second type considers the shape characteristics of the whole region of interest. Contour-based features are Convex, Convex Hull Volume, Convex Hull Volume 3D, Max3D Diameter, Mean Breadth, Orientation, and Surface Area. Area-based features are Compactness1, Compactness2, Mass, Number of Voxel, Roundness, Spherical Disproportion, Sphericity, Surface Area Density, and Volume.

Results and discussion

In recent years, with the rapid development of computing technology, more and more studies have been conducted on the clinical application of image histology as a biomarker for disease detection and therapeutic evaluation. However, before

applying imaging histological features in clinical practice, some problems must be solved, that is, the standardization and robustness of image features. The most important challenge is the repeatability of imaging histological features. In studies of imaging histology, not all image features are recommended. The reason is that some features may lack stability. For example, if the influence of differences in tumor segmentation (differences between manual and semi-automatic segmentation) on image histology features is neglected, phenotypic information of tumors may not be correctly interpreted and the findings will not be repetitive. Therefore, in order to provide consistent and unbiased image biomarkers, it is very important to quantify image histological features repeatedly and objectively.

Image histology often obtains hundreds of image features in clinical research. Potential feature redundancy is another important challenge in image histology research. Redundant image features may increase the complexity of image histology research. Therefore, a set of non-redundant image group biomarkers must be obtained to minimize the over-weighting of redundant image features.

Few studies have evaluated the reproducibility of HCC quantitative CT image features in terms of tumor segmentation methods. The impact of different segmentation of liver tumors on the reproducibility and redundancy of histological features is quantitatively analyzed.

Intra-class correlation coefficient (ICC): ICC is introduced to quantify the repeatability of features. ICC describes the similarity of different measurements within the same group to evaluate the reproducibility of histological features extracted from tumor regions by three segmentation methods. ICC is a descriptive statistics proposed by McGraw and along, whose value is somewhere between 0 and 1, 0 means no repeatability, and 1 means excellent repeatability.

In order to evaluate the influence of sketching differences among different experimenters on feature repeatability, the variance estimates of ICC (A, 1) and bidirectional mixed effect models are calculated.

$$ICC(A,1) = \frac{MS_R - MS_E}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)} \quad (2)$$

ICC (C, 1) and one-way ANOVA (analysis of variance) are calculated to evaluate the effect of different sketches on feature repeatability in the same experimenter.

$$ICC(C,1) = \frac{MS_R - MS_W}{MS_R + (k-1)MS_W} \quad (3)$$

In the above two formulas, MS_R is mean square for rows (observations, fixed factor), MS_W is mean square for residual sources of variance, MSE is mean square error and MSS is mean square for columns (experimenter, random factor). The number of experimenters and the number of observations are expressed by k and n, respectively.

At the same time, as a reliability evaluation parameter, ICC has a variety of evaluation criteria. The criteria proposed by Cicchetti are used to evaluate the repeatability of image features obtained by different segmentation, as follows [11]:

ICC < 0.40, poor repeatability (Poor);
0.40 ≤ ICC < 0. 60, fair repeatability (Fair);
0.60 ≤ ICC < 0.75, good repeatability (Good);
0.75 ≤ ICC ≤ 1, excellent repeatability (Excellent).

Wilcoxon rank-sum Tests or Paired t-test are used to evaluate the difference of ICC values between manual sketching and two semi-automatic segmentation images. When the data conform to normal distribution, Paired t-test is used; when the data does not conform to normal distribution, Wilcoxon rank-sum Tests are used. $P < 0.05$ indicates that the difference has statistical significance.

Repeatability of imaging histological features: ICC of 71 imaging histological features are calculated, including first-order statistical features (intensity histogram features), shape features and texture features. Compared with the features extracted from manual sketch (ICC = 0.80 ± 0.21) and Graph Cut segmentation (ICC = 0.82 ± 0.24), the features extracted from Grow Cut segmentation have higher repeatability (ICC = 0.87 ± 0.19, $P < 0.001$). At the same time, the ICC value of Graph Cut-based imaging histological features is also higher than that of manually sketched, and the difference is statistically significant ($P = 0.036$). For ICC with manual, Graph Cut and Grow Cut imaging features, the average confidence intervals are (0.608, 0.954), (0.774, 0.938) and (0.752, 0.967), respectively.

Among the 71 image histological features, 53 features of Grow Cut have higher ICC values than manual segmentation, and 47 features of Graph Cut have higher ICC values ($P < 0.001$). The ICC value of Grow Cut with 52 features is higher than that of Graph Cut.

Graph Cut (ICC = 0.76 ± 0.26) has no significant difference ($P = 0.332$) compared with intensity histogram features and manual methods (ICC = 0.77 ± 0.29), and Grow Cut (ICC = 0.90 ± 0.12) has obvious high repeatability ($P < 0.001$). For the GLOM features of texture feature categories, Graph Cut (ICC = 0.89 ± 0.11) and Grow Cut (ICC = 0.91 ± 0.12) are more repeatable than manual methods (ICC = 0.77 ± 0.23), respectively ($P = 0.010$, $P = 0.004$). In addition, for GLRLM, NGTDM features and shape features of texture feature categories, there is no significant difference in the repeatability of the three segmentation methods.

All features are classified into four categories according to ICC values. The reproducibility of manual, Graph Cut and Grow Cut is 73, 77 and 81%, respectively, as shown in Fig. 1.

Hierarchical clustering: Hierarchical clustering is to cluster layer by layer, merge and cluster small classes from bottom to top. Specifically, each time two classes with the closest distance are found, then they are merged into a large class. Finally, all the small classes are merged into one class. The whole process is to construct a clustering tree graph, as shown in Fig. 2.

Selection of the best number of clusters: In hierarchical clustering, the determination of the best number of clusters is very important. The commonly used methods to determine the best number of clusters are as follows:

- First, determine by appropriate threshold;
- Second, determine the number of clusters intuitively according to the distribution map of data points.
- Third, determine the number of clusters according to statistics:

– R^2 statistics: The larger the proportion of sum of squares of deviations among different classes is, the smaller the proportion of sum of squares of deviations within classes is. This shows that the clustering results are good. R^2 statistics is the proportion of sum of squares of deviations between classes and sum of squares of deviations within classes;
– Semi-biased R^2 statistics: The difference between k + 1 hierarchical clustering R^2 statistics and k-times merged R^2 statistics;
– False F statistics;
– False t^2 statistics.

R^2 statistics is used to determine the best number of clusters in hierarchical clustering tree graph. The formula is as follows:

$$R^2 = 1 - \frac{P_G}{W} \tag{4}$$

In Formula (4), P_G indicates the sum of squares of intra-class deviations and W represents the sum of squares of total deviations. R^2 is large, which shows that when it is clustered into G class, the squares of deviations within the class are smaller, that is to say, the classification of G class is appropriate. But the more classifications there are, the smaller the sum of squares of deviations in each class will be, and the bigger R^2 will be. Therefore, it can only choose the appropriate G to make R^2 big enough. While G is small, with the increase of G, the increase of R^2 will be small.

Conclusion

Manual tumor segmentation and Graph Cut, Grow Cut semi-automatic tumor segmentation methods are used to segment HCC tumors. Graph Cut is an interactive tumor segmentation method, which achieves tumor segmentation

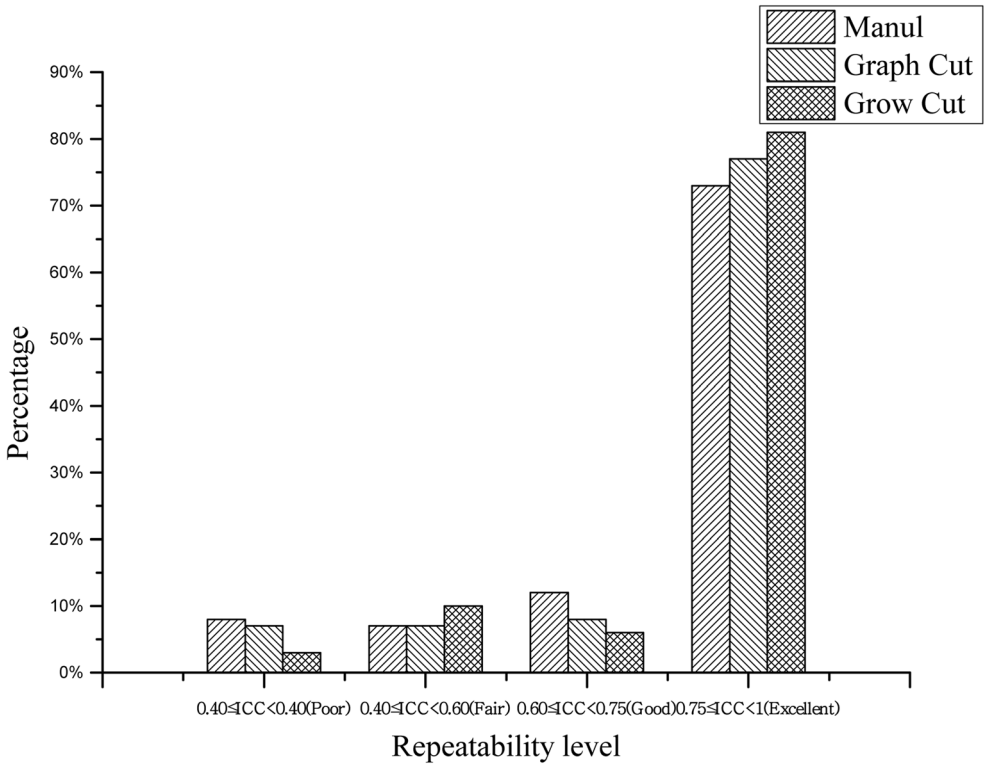

Fig. 1 Percentage of features extracted by three segmentation methods under different ICC

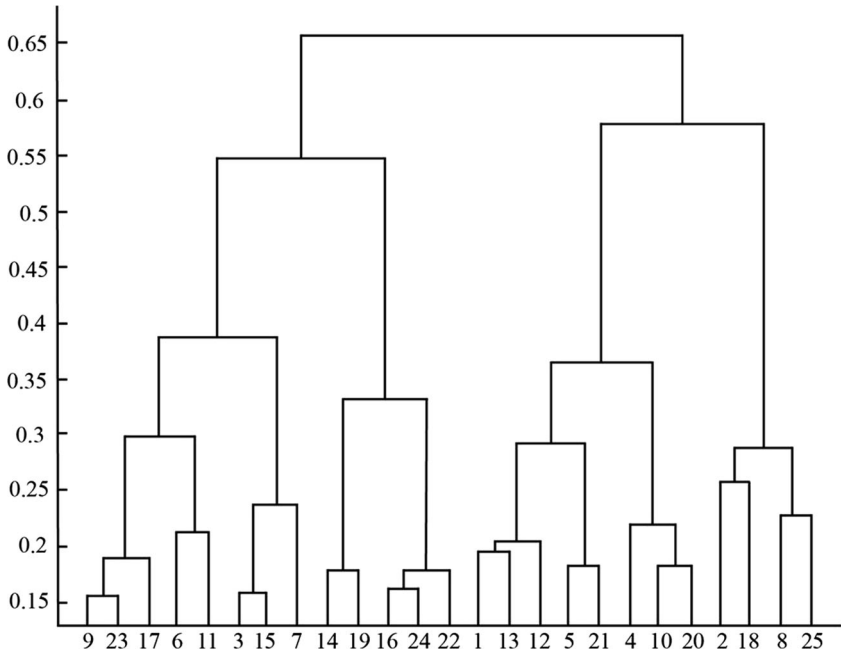

Fig. 2 Hierarchical clustering tree diagram

based on global optimization of energy function. Grow Cut is a segmentation method based on the growth of competitive regions. The principle and initialization of two methods are different, and the influence on image histology cannot be ignored.

It is found that two semi-automatic methods provided by 3D Slice software can extract more stable and repeatable imaging histological features. Compared with manual method, Graph Cut ($P = 0.001$) and Grow Cut ($P < 0.001$) can obtain more reproducible GLOM. The repeatability shows that Grow Cut method has high reproducibility between the two methods. It shows that different initialization states can extract more stable and repeatable imaging histological features. The results show that semi-automatic segmentation can extract more stable and repeatable imaging histological features, especially texture features. At the same time, because of different principles, different segmentation methods can obtain different regions of tumors. Grow Cut semi-automatic tumor segmentation method can provide more stable and repeatable imaging histological features in HCC imaging histology research.

In summary, this study reveals the repeatability difference of histological features extracted from different tumor regions by a variety of segmentation methods. In HCC CT images, the repeatability and redundancy of histological features largely depend on the method of tumor segmentation. Semi-automatic segmentation method can improve the repeatability of image features, and hierarchical clustering can provide stable image feature clustering and reduce feature redundancy. In addition, in order to ensure the segmentation accuracy and avoid the impact of segmentation to the greatest extent, appropriate segmentation algorithms should be considered in different tumor researches. Most importantly, the most stable and unified image features should be selected in the clinical application of image histology.

Compliance with ethical standards

Conflict of interest Author Sen Zhao declares that he has no conflict of interest. Author Wenyan Ren declares that he has no conflict of interest. Author Yan Zhuang declares that he has no conflict of interest. Author Zhixue Wang declares that he has no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

This article does not contain any studies with animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Clark, T., Maximin, S., Meier, J., Pokharel, S., and Bhargava, P., Hepatocellular carcinoma: Review of epidemiology, screening, imaging diagnosis, response assessment, and treatment. Curr. Probl. Diagn. Radiol. 44(6):479–486, 2015.
2. Chen, C. Y., Chen, J., Xia, C. C., Huang, Z. X., and Song, B., T1 mapping combined with Gd-EOB-DTPA-enhanced magnetic resonance imaging in predicting the pathologic grading of hepatocellular carcinoma. J. Biol. Regul. Homeost. Agents 31(4):1029–1036, 2017.

3. Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D., and Bray, F., Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 136(5):359–386, 2015.
4. Scholz, M., Konen, W., Tombrock, S., Fricke, B., Adams, L., von Düring, M., Hentsch, A., Heuser, L., and Harders, A. G., Development of an endoscopic navigation system based on digital image processing. *Computer Aided Surgery* 3(3):134–143, 2015.
5. Sayed, G. I., Hassanien, A. E., and Schaefer, G., An automated computer-aided diagnosis system for abdominal CT liver images. *Procedia Computer Science* 90:68–73, 2016.
6. Kadoury, S., Vorontsov, E., and Tang, A., Metastatic liver tumour segmentation from discriminant Grassmannian manifolds. *Phys. Med. Biol.* 60(16):6459–6478, 2015.
7. Lee, D. H., and Lee, J., Primary malignant tumours in the non-cirrhotic liver. *Eur. J. Radiol.* 95:349, 2017.
8. Petré, C., Ni, Y., Marcha, G., Yu, J., Wevers, M., Lauffer, R. B., and Baert, A. L., Detection and characterization of primary liver cancer in rats by MS-264-enhanced MRI. *Magnetic Resonance in Medicine Official Journal of the Society of Magnetic Resonance in Medicine* 35(4):532–539, 2015.
9. Choi, J. Y., Lee, J. M., and Sirlin, C. B., CT and MR imaging diagnosis and staging of hepatocellular carcinoma: Part II. Extracellular agents, hepatobiliary agents, and ancillary imaging features[J]. *Radiology* 273(1):30–50, 2014.
10. Fujita, N., Asayama, Y., Nishie, A., Ishigami, K., Ushijima, Y., Takayama, Y., Okamoto, D., Moirta, K., Shirabe, K., Aishima, S., Wang, H., Oda, Y., and Honda, H., Mass-forming intrahepatic cholangiocarcinoma: Enhancement patterns in the arterial phase of dynamic hepatic CT - correlation with clinicopathological findings. *Eur. Radiol.* 27(2):1–9, 2016.
11. Brown, K. T., Do, R. K., Gonen, M. et al., Randomized trial of hepatic artery embolization for hepatocellular carcinoma using doxorubicin-eluting microspheres compared with embolization with microspheres alone[J]. *J. Clin. Oncol.* 34(17):2046, 2016.