ImageCLEF Liver CT Image Annotation Task 2014

Neda B.Marvasti[1], Nadin Kökciyan[2], Rüştü Türkay[3], Abdülkadir Yazıcı[2], Pınar Yolum[2], Suzan Üsküdarlı[2], and Burak Acar[1]

[1] Boğaziçi University, Electrical and Electronics Department, Istanbul, Turkey
[2] Boğaziçi University, Computer Engineering Department, Istanbul, Turkey
[3] Istanbul University, Faculty of Medicine, Istanbul, Turkey
{neda.barzegarmarvasti@boun.edu.tr,acarbu@boun.edu.tr}

Abstract. The first Liver CT annotation challenge was organized during the 2014 Image-CLEF workshop held in Sheffield, UK. This challenge entailed the annotation of Liver CT scans to generate structured reports. This paper describes the motivations for this task, the training and test datasets, the evaluation methods, and discusses the approaches of the participating groups. *abstract* environment.

Keywords: ImageCLEF, Liver CT annotation task, Automatic annotation

1 INTRODUCTION

ImageCLEF [1] was part of the Cross Language Evaluation Forum (CLEF) 2014 consisting of four main tasks: Robot Vision, Image Annotation, Liver CT Annotation, and Domain Adaptation. It was the first time that the automatic annotation of Liver CT images was provided as a challenge.

The purpose of the Liver CT annotation task was to automatically generate structured reports with the use of computer generated features of liver CT volumes. Structured reports are highly valuable in medical contexts due to the processing opportunities they provide, such as reporting, image retrieval, and computer-aided diagnosis systems. However, structured reports are cumbersome and time consuming to create. Furthermore, their creation requires domain expertise who is time constrained. Consequently, such structured medical reports are often not found or are incomplete in practice. This challenge was designed to aid the generation of structured reports.

The datasets provided for this challenge consisted of 50 training and 10 test datasets. Participants were asked to answer a fixed set of multiple-choice questions about livers. The questions were automatically generated from an open-source ontology of liver for radiology (onlira) [2]. The answers to the questions describe the properties of the liver, the hepatic vasculature of the liver, and a specific lesion within the liver. During this task, the user is presented with the following training data: (1) data from a CT scan, (2) a liver mask, (3) a

volume-of-interest that highlights the selected lesion, and (4) a rich set of imaging observations. The imaging observations are ONLIRA based annotations that were manually entered by radiologists. Participants were permitted to extract their own image features from the CT data and use them. The results were evaluated in terms of the completeness and accuracy of the generated report.

The rest of the paper is organized as follows, Section 2 gives a detailed description of the task and introduces the participants. Section 3 presents the main results of the task and the results of the participants, and Section 4 concludes the paper.

2 The Liver CT Annotation Challenge

This section describes the task and introduces the participants.

2.1 Task Definition and Datasets

The Liver CT annotation task is proposed towards the generation of structured reports describing the semantic features of the liver, its vascularity, and the types of lesions in the liver. The goal of proposing this task is to develop automated mechanisms to assist in the difficult and practically infeasible task of annotating medical records.

The training dataset includes 50 cases, each consisting of:

- a cropped CT image of the liver – a 3D matrix with the same size as cropped CT image,
- a liver mask that specifies the part corresponding to the liver – a 3D matrix indicating the liver areas with a 1 and nonliver areas with a 0,
- a bounding box (ROI) corresponding to the region of the selected lesion within the liver – as a vector of 6 numbers corresponding to the coordinates of two opposite corners,
- a set of 60 computer generated (CoG) features obtained from an interactive segmentation software – a 60×4 array, and
- A set of 73 user expressed features (UsE) manually entered by a radiologist and stored in a 73×6 array.

In training dataset 50 .mat files, each consisting of all the above data were given to the participants. The format of the test dataset is the same except that the UsE features are missing, which the participants were expected to predict. The participants were allowed to extract and use their own image features. It is important to note that the resolution of CT images may vary ($x : 190 - 308$ pixels, $y : 213 - 238$ pixels, and $z : 41 - 588$ slices). The spacing may also vary in the range of ($x, y : 0.674 - 1.007$ mm, $slice : 0.399 - 2.5$ mm).

Computer Generated Features For each case, there is a set of 60 CoG image descriptors. Table 4 provides the list of all CoG features for a case. Some of them have only one value and the rest are vectors with different dimension. For example, the size of "HistogramOfAllLesions" is67, while the size of "LiverVolume" is 1. The total dimension of all features is 454. The CoG features were extracted after interactively marking the liver, vessels, and lesions on a CT image. The CoG features describe the characteristics of the liver, vessels, and lesions. Lesion descriptors are categorized into five types: geometric, locational, gray-scale, boundary, and texture features. The reader is referred to [3] for the details of CoG features. In both training and test dataset, CoG features are shown in a 60 array where each columns is stands for:

Column	Feature	Type
1	group	string
1	name	string
3	type	string
4	value	type of the feature

User Expressed Features Imaging observations of a radiologist for the liver domain are represented with ONLIRA(Ontology of the Liver for Radiologists). A web based data collection application, called CaReRa-Web [4].

For each case, there are 73 user expressed (UsE) features represented in a 73×6 cell array. These features clinically characterize the liver, hepatic vascularity, and liver's lesions. In the training dataset, the UsE features are manually entered by an expert radiologist. Every UsE feature corresponds to a question answered by a radiologist. Some UsE features may take on more than one value. Such features are represented with a multi-selection answers.

In the test phase, the participants were expected to predict the UsE features. The format of the 73×6 UsE data is:

Column	Annotation Features	Type
1	Group	string
2	Concept	string
3	Properties	string
4	Indices	bar separated list of integers
5	Values	bar separated list of strings
6	"Free text" related to value	Text

The "Group" and "Concept" are the ONLIRA-based concepts. Each concept may have several properties. Each property may have multiple values whose indices and meaning are given in "Indices" and "Values" columns, respectively. Properties deemed irrelevant are marked as NA by the radiologist. UsE features are grouped as: Liver, Vessel, General and Lesion. Table 5 lists every group and its corresponding concepts, properties, possible values and their assigned indices.

[4] The CaReRa-Web is a tool that can accessed at `https://vavlab.ee.boun.edu.tr:5904/CareraWeb2`. It is available for academic use from the CaReRaproject (Case Retrieval in Radiological Databases) website `http://www.vavlab.ee.boun.edu.tr`

2.2 Evaluation methodology

The evaluation is performed on the basis of the completeness and accuracy of the predicted annotations with reference to the manual annotations of the test dataset. Completeness is defined as the number of predicted features divided by total number of features, while accuracy is the number of correct predicted features divided by total number of predicted features.

For answers that allow multiple values to a question, the correct prediction of a single feature is considered as a correct annotation.

$$Completeness = \frac{number\ of\ predicted\ UsE\ \ features}{Total\ number\ of\ UsE\ \ features} \tag{1}$$

$$Accuracy = \frac{number\ of\ correctly\ predicted\ UsE\ \ features}{Number\ of\ predicted\ UsE\ \ features} \tag{2}$$

$$TotalScore = \sqrt{Completeness \times Accuracy} \tag{3}$$

2.3 Participation

Among 20 groups, which registered for the task and signed the license agreement to access the datasets, only 3 of them submitted at least one run. The number of runs per group was limited to ten. Tables 1 describes these runs.

Table 1: ImageCLEF Liver CT Image Annotation Task 2014 participants who performed at least one run.

Group name	**Affiliation**	**Num of runs**
BMET	School of Information Technologies, University of Sydney, Australia	8
CASMIP	The Hebrew University of Jerusalem, Israel	1
piLabVAVlab	Boğaziçi University, Turkey	1

3 Results

The groups that submitted their results based their prediction on classifiers, image retrieval, and generalized coupled tensor factorization (GCTF).

The BMET group, achieved the best results using the image retrieval techniques with total score of %94.7. Classifier-based methods were used by both BMET and CASMIP groups. Only piLabVAVlab used a GCTF method. Table 2 shows the completeness, accuracy and total score achieved by each run on the test dataset.

The BMET group [4] submitted 8 runs, of which 4 of them used a classifier-based approach and the remaining used an image retrieval algorithm. They used two different feature sets: the prepared CoG features from the database and a

bag of visual words (BoVW). In the classifier-based approach, they used two-stage classification, where each stage consists of a bank of several support vector machines (SVM), which is used for each UsE feature. A two-stage classification is proposed to solve the unbalanced training dataset. For each UsE feature, the first stage is composed of the 1-vs-all classifiers and the second stage is consisted of the 1-vs-1 classifiers. Second classifier is activated only if the result of the first step is more than one label. In the second stage they run 1-vs-1 classifier for the set of labels resulting from the first step and a majority voting scheme is used to select the final answer. In their first and second runs, they used linear kernels while in their third and fourth runs, they employed radial basis function (RBF) kernel. They examined these two kernels with both sets of features.

In the image retrieval based approach, they used the most similar training images to select the UsE features for the test image. Similarity is calculated by computing the Euclidean distance between image feature vectors. Finally, they applied a weighted voting scheme to select the label assigned to each UsE features using the "n" most similar images to the test image, where $n = 10$ in this scenario. Basically, this algorithm votes images more similarity to the test image with higher values. Results of this approach with two different sets of features are seen in $5th$ and $6th$ runs. In $7th$ and $8h$ runs, they applied a sequential feature selection method to use the most discriminating features for each question during the similarity calculation, in order to use the most suitable one. As mentioned above, BMET group used two kernels for SVM classification, however, there is no significant performance difference in the results, which the participants attribute to the unbalanced training dataset. Their classification methods performed best when they employed their expanded feature set. Their retrieval method performed best when the given CoG features were employed. This observation suggests that the nature of feature sets are important for utilizing different methods.

CASMIP group [5] submitted one run to the task, which achieved the second best performance. They tried four different classifiers in the learning phase: linear discriminant analysis (LDA), logistic regression (LR), K-nearest neighbors (KNN), and finally SVM to predict UsE features. An exhaustive search of every combination of image features is done using leave-one-out cross validation method on training data for every UsE feature and classifier. As the result, for each UsE feature the best classifier and its related features are learned. They used only a certain part of provided CoG features, which was achieved by ignoring 21 high-dimension features, i.e. they ignored features with dimensionally more than one. Instead, 9 additional features have been added to individual lesion features extracted in the lesion ROI describing the gray level features of liver, lesion, and boundary of lesion properties. The learning step was performed using all UsE features of the training dataset except cluster size, lobe and segments, which were obtained directly from image features. Python scikit-Learn Machine learning toolbox was used for implementing each classifier with the default parameters. As the result, for most of the UsE features they got same performance using any classifier and any combination of image features. Hence they assigned

any classifier and all image features for them. For 6 of the UsE features that describe the density, contrast and location of lesions, one of the LDA or KNN classifiers was chosen along with their selected features.

piLabVAVlab group [6] considered the dataset as heterogeneous data and GCTF approach was applied to predict UsE features. They considered both KL-divergence and Euclidean-distance-based cost functions as well as the coupled matrix factorization models using GCTF framework. They tried to predict approximately half of the UsE features. In order to achieve this, UsE features with only 4 indices whose values vary from 0 to 3 were considered as the first group and UsE features with binary indices were considered as the second groups. The reason for this was that the threshold selection needed to be specified for each type of question. Thus, they considered questions with similar answer ranges in a study and ignore question with varied answer ranges. Basically, they provide three matrices for 50 training and 10 test cases:

- X_1: A 60×21 matrix (UsE features of first group).
- X_2: A 60×13 matrix (UsE features of second group).
- Z_1: A 60×447 matrix (all CoG features).

They estimated the latent matrices: Z_2 and Z_3 by using coupled matrix factorization models according to the following formula:

$$X_1 = Z_1 * Z_2 \tag{4}$$
$$X_2 = Z_1 * Z_3 \tag{5}$$

The UsE features of the 10 test cases are predicted with Z_2 and Z_3. Since the predicted values are not discrete, a binary thresholding method has been proposed to extract the labels of UsE features.

This group submitted one run during the submission period, which had the accuracy of %45. However, after the submission deadline, they claimed that they had improved their thresholding method and requested that we evaluate their new results (see Table 2 run2 and run3).

Among 73 UsE features, 7 of them were excluded from the evaluation because of their unbounded labels (numeric continuous values). The BMET group achieved the highest scores with completeness of %98 (See Table 2. In terms of accuracy, BMET group has also attained the best performance by using an image retrieval method. In terms of classifier-based methods, BMET and CASMIP groups both obtained the total score of %93.

Table 3 compares the results of different runs in predicting different groups of UsE features. We divide UsE features into 5 groups: liver, vessels and three lesion groups with area, lesion and component concepts. Results show that all the groups have completed the vessel UsE features with high accuracy. The BMET and CASMIP groups completed liver features in full with accuracy more than %80. None of the groups can completely annotate the area related concepts of lesions. Components related concepts of lesion are completed fully and annotated

Table 2: Results of the runs of Liver CT annotation task. CoG: just given CoG features are used. CoG+: user generated features are added to given CoGfeatures.

Group name	Run	Completeness	Accuracy	Total Score	method used	feature used
BMET	run1	0.98	0.89	0.935	SVM-linear	CoG
BMET	run2	0.98	0.90	0.939	SVM-linear	CoG+
BMET	run3	0.98	0.89	0.933	SVM-RBF	CoG
BMET	run4	0.98	0.90	0.939	SVM-RBF	CoG+
BMET	run5	0.98	0.91	0.947	IR-noFS	CoG
BMET	run6	0.98	0.87	0.927	IR-noFS	CoG+
BMET	run7	0.98	0.91	0.947	IR-FS	CoG
BMET	run8	0.98	0.87	0.926	IR-FS	CoG+
CASMIP	run1	0.95	0.91	0.93	LDA+KNN	CoG+
piLabVAVlab	run1	0.51	0.39	0.45	MF-KL	CoG
piLabVAVlab	run2	0.51	0.89	0.677	MF-EUC	CoG
piLabVAVlab	run3	0.51	0.88	0.676	MF-KL	CoG

with accuracy higher than %72 by both BIMET and CASMIP groups. Lesion related concepts of lesions are annotated completely by only BIMET group with accuracy more than %72.

Table 3: Completeness(complete.) and Accuracy(acc.) for five different groups of UsE features

GroupName	Liver		Vessel		LesionArea		LesionLesion		LesionComponent	
name	complete.	acc.	complete.	acc.	complete.	acc.	complete.	acc.	complete.	acc.
BMET-run1	1.00	0.91	1.00	1.00	0.92	0.78	1.00	0.72	1.00	0.93
BMET-run2	1.00	0.93	1.00	1.00	0.92	0.77	1.00	0.77	1.00	0.94
BMET-run3	1.00	0.93	1.00	1.00	0.92	0.76	1.00	0.72	1.00	0.93
BMET-run4	1.00	0.93	1.00	1.00	0.92	0.77	1.00	0.77	1.00	0.94
BMET-run5	1.00	0.93	1.00	1.00	0.92	0.79	1.00	0.83	1.00	0.94
BMET-run6	1.00	0.80	1.00	1.00	0.92	0.72	1.00	0.79	1.00	0.93
BMET-run7	1.00	0.93	1.00	1.00	0.92	0.79	1.00	0.83	1.00	0.94
BMET-run8	1.00	0.93	1.00	1.00	0.92	0.68	1.00	0.73	1.00	0.92
CASMIP	1.00	0.93	1.00	1.00	0.85	0.81	0.90	0.82	1.00	0.94
piLabVAVlab-run1	0.62	0.77	1.00	0.42	0.46	0.20	0.20	0.00	0.12	0.15
piLabVAVlab-run2	0.62	0.88	1.00	1.00	0.46	0.77	0.20	1.00	0.12	0.15
piLabVAVlab-run3	0.62	0.88	1.00	0.99	0.46	0.77	0.20	1.00	0.12	0.15

4 Conclusion

This was the first time the liver CT annotation task was proposed. We provided liver patient data collected via a hybrid patient information entry system whose

liver characteristics are based on the ONLIRA ontology. The challenge presented to the participants was to predict USE features of patient records, given COG features. As this was the first time for this challenge, it was not surprising that few groups were able to submit their runs for this complex task. out of 20 teams 3 teams submitted at least 1 run. The approaches and results were reviewed and documented in this paper.

The main challenge of the task was due to the unbalanced dataset and participants tried to overcome this issue with different methods. Among all methods image retrieval obtained the best performance. It was observed that feature selection is important for the best performance of the prediction method.

Acknowledgments This work is in part supported by CaReRa Project (TÜBİTAK Project No: 110E264) and Bogazici University B.A.P (Project No: 5324)

References

1. Barbara Caputo, Henning Müller, Jesus Martinez-Gomez, Mauricio Villegas, Burak Acar, Novi Patricia, Neda Marvasti, Suzan Üsküdarlı, Roberto Paredes, Miguel Cazorla, Ismael Garcia-Varea, and Vicente Morell, "ImageCLEF 2014: Overview and analysis of the results," in *CLEF proceedings*, Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2014.
2. N. Kokciyan, R. Turkay, S. Uskudarli, P. Yolum, B. Bakir, and B. Acar, "Semantic description of liver ct images: An ontological approach," 2014.
3. Neda Barzegar Marvasti, Ceyhun Burak Akgül, Burak Acar, Nadin Kökciyan, Suzan Üsküdarlı, Pınar Yolum, Rüstü Türkay, and Barıs Bakır, "Clinical experience sharing by similar case retrieval," in *Proceedings of the 1st ACM international workshop on Multimedia indexing and information retrieval for healthcare*. ACM, 2013, pp. 67–74.
4. Ashnil Kumar, Shane Dyer, Changyang Li, Philip H. W. Leong, and Jinman Kim, "Automatic annotation of liver ct images: the submission of the bmet group to imageclefmed 2014," in *CLEF 2014 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings (CEUR-WS.org)*, September 2014.
5. Assaf B. Spanier and Leo Joskowicz, "Towards content-based image retrieval: From computer generated features to semantic descriptions of liver ct scans," in *CLEF 2014 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings (CEUR-WS.org)*, September 2014.
6. Beyza Ermis and A. Taylan Cemgil, "Liver ct annotation via generalized coupled tensor factorization," in *CLEF 2014 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings (CEUR-WS.org)*, September 2014.

Table 4: List of CoGfeature

Group	Name	Description	#
Liver	LiverVolume	Liver volume(mm)	1
	LiverMean	Liver's mean intensity value	1
	LiverVariance	Liver's variance intensity value	1
Vessel	VesselRatio	|vessel's voxels|/ |liver's voxels|	1
	VesselVolume	Vessel volume (mm)	1
AllLesions	NumberofLesions	Number of lesions in the liver.	1
	MinLesionVolume	Smallest lesion's volume (mm)	1
	MaxLesionVolume	Biggest lesion's volume (mm)	1
	LesionRatio	|lesions' voxels|/ |liver's voxels|	1
	HistogramOfAllLesions	Histogram of all lesions intensity values.	67
	Mean	Lesions' mean intensity	1
	Variance	Lesions' variance intensity	1
	Skewness	Lesions' skewness	1
	Kurtosis	Lesions' kurtosis	1
	Energy	Lesions' energy	1
	Entropy	Lesions' entropy	1
	Smoothness	Lesions' smoothness	1
	Abscissa	Bin of histogram peak	1
	Threshold	# voxels with intensity values more than liver mean.	1
	HaarWaveletCoeff	3 level Haar wavelet coeff. of histogram.	8
Lesion	Histogram	Histogram of lesion intensity value.	67
	Mean	Lesion's mean intensity	1
	Variance	Lesion's variance intensity	1
	Skewness	Lesion's skewness	1
	Kurtosis	Lesion's kurtosis	1
	Energy	Lesion's energy	1
	Entropy	Lesion's entropy	1
	Smoothness	Lesion's smoothness	1
	Abscissa	Bin of histogram peak.	1
	Threshold	# voxels with intensity values more than liver mean.	1
	HaarWaveletCoeff	3 level Haar wavelet coeff. of histogram.	8
	AnatomicalLocation	5 anatomically identifiable locations. (mm)	5
	Lesion2VesselMinDistance	Min. dist. bet. lesion's center and nearest vessel (mm).	1
	Lesion2VesselTouchRatio	Touch ratio between lesion and vessels.	1
	VesselTotalRatio	Ratio of vessel-volume to liver-volume.	1
	VesselLesionRatio	Ratio of vessel-volume to lesion-volume.	1
	Volume	Volume of lesion in mm.	1
	SurfaceArea	SurfaceArea of lesion in mm.	1
	MaxExtent	Max lesion radios in mm.	1
	AspectRatio	Ratio of max lesion to min lesion radios(mm).	1
	Sphericity	Sphericity of lesion in mm.	1
	Compactness	Compactness of lesion in mm.	1
	Convexity	Convexity of lesion in mm.	1
	Solidity	Solidity of lesion in mm.	1
	FourierDescriptors	A 20D vector of the lesion's Fourier descriptors.	20
	BoundaryScaleHistogram	Histogram of boundary scale values.	33
	BoundaryWindowHistogram	Histogram of boundary window values.	33
	HaralickEnergy	Lesion's Haaralicks Energy.	4
	HaralickEntropy	Lesion's Haaralicks Entropy.	4
	HaralickInverseDiffMoment	Lesion's Haaralicks Inverse Difference Moment.	4
	HaralickInteria	Lesion's Haaralicks Interia.	4
	HaralickClusterShade	Lesion's Haaralicks Cluster Shade.	4
	HaralickClusterProminence	Lesion's Haaralicks Cluster Prominence.	4
	HuMoments	Lesion's 3D hu moments.	3
	TamuraCoarsenessHistogram	Lesion's Tamura Coarseness.	7
	TamuraContrastHistogram	Lesion's Tamura Contrast.	19
	TamuraDir1Histogram	Hist. of lesion's Tamura x directionality.	19
	TamuraDir2Histogram	Hist. of lesion's Tamura y directionality.	19
	TamuraDir3Histogram	Hist. of lesion's Tamura z directionality.	19
	GaborEnergy	Lesion's Gabor energies in 4 scales and 16 directions.	64

Table 5: List of UsE features

Group	Concept	Properties	Possible values(assigned indices)
Liver	Liver	Liver Placement	downward displacement(0), normal placement(1), leftward displacement(2), upward displacement(3), other(4)
		Liver Contour	irregular(0), lobulated(1), nodular(2), regular(3), other(4)
		Liver Size Change	decreased(0), increased(1), normal(2), other(3)
		Liver Craniocaudal Dimension(mm)	The amount change in size of liver(mm)
		Density Type	heterogeneous(0), homogeneous(1), other(2)
		Density Change	decreased(0), increased(1), normal(2), other(3)
	Right Lobe	Right Lobe Craniocaudal Dimension(mm)	The amount change in size of right lobe(mm)
		Right Lobe Size Change	decreased(0), increased(1), normal(2), other(3)
	Left Lobe	Left Lobe Craniocaudal Dimension(mm)	The amount change in size of left lobe(mm)
		Left Lobe Size Change	decreased(0), increased(1), normal(2), other(3)
	Caudate Lobe	Caudate Lobe Craniocaudal Dimension(mm)	The amount change in size of caudate lobe(mm)
		Caudate Lobe Size Change	decreased(0), increased(1), normal(2), other(3)
Vessel	Hepatic Artery	Hepatic Artery Lumen Diameter	decreased(0), increased(1), normal(2), other(3)
		Hepatic Artery Lumen Type	obliterated(0), open(1), partially obliterated(2), other(3)
	Hepatic Portal Vein	Hepatic Portal V. Lumen Diam.	decreased(0), increased(1), normal(2), other(3)
		Hepatic Portal V. Lumen Type	obliterated(0), open(1), partially obliterated(2), other(3)
		is Cavernous Transformation Observed?(Hepatic Portal Vein)	NA(-1),True(1),False(0),NA(-1)
	Left Portal Vein	Left Portal V. Lumen Diam.	decreased(0), increased(1), normal(2), other(3)
		Left Portal V. Lumen Type	obliterated(0), open(1), partially obliterated(2), other(3)
		is Cavernous Transformation Observed?(Left Portal Vein)	NA(-1),True(1),False(0),NA(-1)
	Right Portal Vein	Right Portal V. Lumen Diam.	decreased(0), increased(1), normal(2), other(3)
		Right Portal V. Lumen Type	obliterated(0), open(1), partially obliterated(2), other(3)
		is Cavernous Transformation Observed?(Right Portal Vein)	NA(-1),True(1),False(0),NA(-1)
	Hepatic Vein	Hepatic V. Lumen Diam.	decreased(0), increased(1), normal(2), other(3)
		Hepatic V. Lumen Type	obliterated(0), open(1), partially obliterated(2), other(3)

Group	Concept	Properties	Possible values(assigned indices)
Vessel	Left Hepatic Vein	Left Hepatic V. Lumen Diam.	decreased(0), increased(1), normal(2), other(3)
		Left Hepatic V. Lumen Type	obliterated(0), open(1), partially obliterated(2), other(3)
	Middle Hepatic Vein	Middle Hepatic V. Lumen Diam.	decreased(0), increased(1), normal(2), other(3)
		Middle Hepatic V. Lumen Type	obliterated(0), open(1), partially obliterated(2), other(3)
	Right Hepatic Vein	Right Hepatic V. Lumen Diam.	decreased(0), increased(1), normal(2), other(3)
		Right Hepatic V. Lumen Type	obliterated(0), open(1), partially obliterated(2), other(3)
General	Patient	Diagnosis	Diagnosis of given image using ICD10 codes (bar separated) and in the free text MD's comments are written (bar separated).
Lesion	Lesion	Cluster Size	1(1), 2(2), 3(3), 4(4), 5(5), multiple(6) For simple cases this value shows number of lesions inside the ROI, but in case of having more than one lesions of a certain type, the biggest lesion is annotated as a sample of that cluster and number of lesions with same properties is written here
		Contrast Uptake	NA(-1), dense(0), heterogeneous(1), homogeneous(2), minimal(3), moderate(4), other(5)
		Contrast Pattern	NA(-1), central(0), early uptake then wash out(1), fixing contrast in late phase(2), heterogeneous(3), homogeneous(4), peripheric(5), peripheric nodular(6), spokes wheel(7), undecided(8), other(9)
		Lesion Composition	SolidCycsticMix(0), Solid(1), SolidWithCystic(2), PureSolid(3), PredominantSolid(4), Cystic(5), PureCystic(6), PredominantCystic(7), CysticWithSolidComponent(8), CysticWithDebris(9), Abcess(10)
		is Leveling Observed?	True(1),False(0)
		Leveling Type	NA(-1), fluid fluid(0), fluid gas(1), fluid solid(2), gas solid(3), other(4)
		is Debris observed?	True(1),False(0),NA(-1)
		Debris Location	NA(-1), floating inside(0), located on dependent position(1),other(2)
		is Close to Vein	NA(-1), HepaticArtery(0), HepaticPortalVein(1), RightPortalVein(2), LeftPortalVein(3), HepaticVein(4), RightHepaticVein(5), MiddleHepaticVein(6), LeftHepaticVein(7), VenacavaInferior(8), PosteriorBranchOfRightPortalVein(9), AnteriorBranchOfRightPortalVein(10), other(11)
		Vasculature Proximity	NA(-1), adjacent(0), adjunct to contact(1), bended(2), circumscribed(3), invaded(4), other(5)
	Area	Lobe	LeftLobe(0), CaudateLobe(1), RightLobe(2)
		Segment	SegmentI(1), SegmentII(2), SegmentIII(3), SegmentIV(4), SegmentV(5), SegmentVI(6), SegmentVII(7), SegmentVIII(8)
		width	a number in mm which represents width of the lesion
		height	a number in mm which represents heigth of the lesion
		is Gallbladder Adjacent?	True(1),False(0)
		is Peripherical Localized?	True(1),False(0)
		is Subcapsular Localized?	True(1),False(0)
		is Central Localized	True(1),False(0)

Group	Concept	Properties	Possible values(assigned indices)
,	Area	Margin Type	geographical(0), ill defined(1), irregular(2), lobular(3), serpiginious(4), spiculative(5), well defined(6), other(7)
		Shape	band(0), fusiform(1), irregular(2), linear(3), nodular(4), ovoid(5), round(6), serpiginious(7), other(8)
		is Contrasted	True(1),False(0),NA(-1)
		is Calcified? (Area)	True(1),False(0),NA(-1)
		Area Calcification Type	NA(-1), coarse(0), focal(1), millimetric-fine(2), punctate(3), scattered(4), other(5)
		Density	NA(-1), hyperdense(0), hypodense(1), isodense(2), other(3)
		Density Type	NA(-1), heterogeneous(0), homogeneous(1), other(2)
	Capsule	is Calcified? (Capsule)	True(1),False(0),NA(-1)
		Capsule Calcification Type	NA(-1), coarse(0), focal(1), millimetric-fine(2), punctate(3), scattered(4), other(5)
	Polyp	is Calcified? (Polyp)	True(1),False(0),NA(-1)
		Polyp Calcification Type	NA(-1), coarse(0), focal(1), millimetric-fine(2), punctate(3), scattered(4), other(5)
	Pseudocapsule	is Calcified?(Pseudocapsule)	True(1),False(0),NA(-1)
		Pseudocapsule Calc. Type	NA(-1), coarse(0), focal(1), millimetric-fine(2), punctate(3), scattered(4), other(5)
	Septa	is Calcified? (Septa)	True(1),False(0),NA(-1)
		Septa Calcification Type	NA(-1), coarse(0), focal(1), millimetric-fine(2), punctate(3), scattered(4), other(5)
		Diameter Type	NA(-1), complete(0), incomplete(1), other(2)
		Thickness	NA(-1), thick(0), thin(1), other(2)
	Solid Component	is Calcified? (Solid Component)	True(1),False(0),NA(-1)
		Solid Component Calcification Type	NA(-1), coarse(0), focal(1), millimetric-fine(2), punctate(3), scattered(4), other(5)
	Wall	is Calcified? (Wall)	True(1),False(0),NA(-1)
		Wall Calcification Type	NA(-1), coarse(0), focal(1), millimetric-fine(2), punctate(3), scattered(4), other(5)
		Wall Type	NA(-1), thick(0), thin(1), other(2)
		is Contrasted?(Wall)	True(1),False(0),NA(-1)