

Synthesis of a Neural Network Classifier for Hepatocellular Carcinoma Grading Based on Triphasic CT Images

Vitoantonio Bevilacqua¹(✉), Leonarda Carnimeo¹, Antonio Brunetti¹,
Andrea De Pace¹, Pietro Galeandro¹, Gianpaolo Francesco Trotta²,
Nicholas Caporusso¹, Francescomaria Marino¹, Vito Alberotanza³,
and Arnaldo Scardapane³

¹ Department of Electrical and Information Engineering, Polytechnic of Bari,
Via Orabona 4, 70125 Bari, Italy

vitoantonio.bevilacqua@poliba.it

² Department of Mechanical and Management Engineering, Polytechnic of Bari,
Via Orabona 4, 70125 Bari, Italy

³ Interdisciplinary Department of Medicine - Section of Diagnostic Imaging,
University of Bari Aldo Moro, Piazza Giulio Cesare 11, 70120 Bari, Italy
arnaldo.scardapane@uniba.it

Abstract. Computer Aided Decision (CAD) systems based on Medical Imaging could support radiologists in grading Hepatocellular carcinoma (HCC) by means of Computed Tomography (CT) images, avoiding that patient undergo any medical invasive procedures such as biopsies. The individuation and characterization of Regions of Interest (ROIs) containing lesions is an important phase that enables an easier classification between two classes of HCCs. Two phases are needed for the individuation of lesioned ROIs: a liver isolation in each CT slice, and a lesion segmentation. Ultimately, all individuated ROIs are described by morphological features and, finally, a feed-forward supervised Artificial Neural Network (ANN) is used to classify them. Testing determined that the ANN topologies found through an evolutionary strategy showed a high generalization on the mean performance indices regardless of applied training, validation and test sets, showing good performances in terms of both accuracy and sensitivity, permitting a correct grading of HCC lesions.

Keywords: Hepatocellular carcinoma · Computed Tomography · Image processing · Artificial Neural Network · Genetic Algorithm

1 Introduction

Hepatocellular carcinoma (HCC) is the fifth most common malignant disease in men and the eighth most common in women. It is the third most common cause of death from cancer, after lung tumor and stomach cancer [1]. HCCs mostly

develop in patients with chronic liver disease caused by viral hepatitis, alcohol abuse and inborn metabolic errors. Eighty percent of all HCCs worldwide occur when the underlying chronic liver disease has reached the cirrhotic stage [2,3]. Moreover, less common causes are Wilson's disease, hereditary hemochromatosis, alpha1-antitrypsin deficiency, primary biliary cirrhosis and autoimmune hepatitis [4]. The nodules associated with liver cirrhosis are histologically divided into six categories according to the classification proposed by the Liver Cancer Study Group of Japan: large regenerative nodules, adenomatous hyperplasia (AH), atypical AH, early HCC, well differentiated HCC, and moderately or poorly differentiated HCC (so-called classical HCC) [5,6]. Namely, HCC differentiation directly impact on patients' prognosis.

Thanks to recent progress, Computed Tomography (CT) and Magnetic resonance Imaging (MRI) are the gold standard for non-invasive evaluation of diffuse and focal diseases of the liver and biliary tract [7,8]. Unlike most cancers, imaging can be used to non-invasively diagnose HCC; treatment, including major surgical options such as hepatic resection and liver transplantation, can be realized without confirmatory biopsy [9,10].

The hallmark diagnostic features of HCC at multiphasic CT or MR imaging are arterial phase hyper-enhancement, followed by portal venous or delayed phase washout appearance; in patients with cirrhosis or other risk factors for HCC, such temporal enhancement pattern provides near 100% specificity for diagnosis of HCC [9]. However, the presence of specific imaging features which may predict the differentiation of HCC is still widely debated.

As HCC tends to be untreatable when it is diagnosed at a late stage, it is necessary to diagnose it in an early stage by means of blood exams using liver cancer biomarkers, or other exams.

According to World Health Organization (WHO), an HCC can be analyzed from both macroscopic and microscopic point of view; at first, the degree of tumor depends on size and presence (or absence) of liver cirrhosis. From a microscopic point of view, a histological classification of tumors includes trabecular (plate-like), pseudoglandular and acinar, compact and scirrhous types [11]. For several years, grading of HCC relied on Edmondson and Steiner system, which divided HCC into four grades - from I to IV - based on histological differentiation [12]. Grade I is the best differentiated form, consisting of small tumor cells arranged in thin trabeculae. Cells in grade II are larger, and they have abnormal nuclei and glandular structures. In grade III, nuclei are larger and more hyperchromatic than grade II cells; also, the cytoplasm is granular and acidophilic, though less than grade II. In grade IV, tumor cells are much less differentiated with hyperchromatic nuclei and loss of trabecular pattern. In particular, most of HCCs appears as grade II or III.

As stated before, thanks to imaging, it is possible to study and determine the stage of liver tumor; one of the most utilized imaging technique in this field is Computed Tomography (CT). Indeed, new technology enables multi-phases acquisition protocols consisting of different acquisitions repeated at specified time intervals.

2 Materials

2.1 Patients

In the time interval between January and August 2016, four patients (men aged 55–78 years, mean \pm standard deviation (SD), 66.3 ± 6.4 years) were subject to liver transplant, partial hepatectomy, or US-guided needle biopsy with pre-operative abdominal CT showing hepatic lesion and the typical HCC pattern. In all the cases histopathological specimen analysis confirmed the diagnosis of overt HCC. On the basis of histopathology, the patients were divided into 2 groups:

- well/moderately differentiated nodule;
- poorly differentiated nodules.

2.2 Computed Tomography Acquisition Protocol

All CT scans were acquired with a 320 slices Scanner (Toshiba Aquilion One) after an automated injection of 1.5 ml/kg of iodinated contrast medium (Iomeprole 400 mgI/ml) through a 16G Needle in antecubital vein at a flow rate of 4 ml/s with the following protocol:

1. **arterial** dominant phase (Fig. 1(a)) acquired 20 s after the aortic peak calculated by a bolus tracking system with a ROI positioned in the abdominal aorta at a trigger density of 150 Hounsfield Units (HU);
2. **portal** phase (Fig. 1(b)) acquired 70 s after contrast injection;
3. **equilibrium** phase (Fig. 1(c)) acquired 180 s after contrast injection.

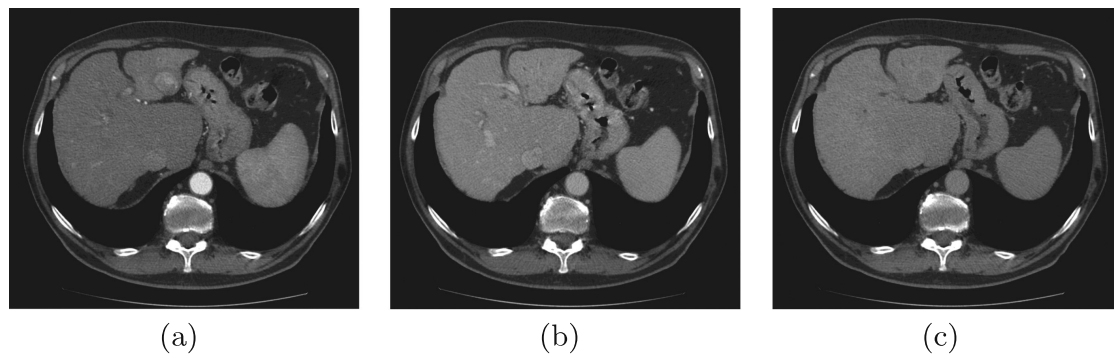


Fig. 1. CT images acquired in the different phases: (a) arterial phase; (b) portal phase; (c) equilibrium phase

For each CT series, we realized a volumetric acquisition using a collimation of 0.6 mm, a pitch value of 1.2, 120 kV, 160–440 mAs. Images were reconstructed with 1 mm as thickness and a reconstruction index of 0.8. Then, they were sent to a post-processing workstation.

3 Methods

Figure 1 shows that the lesion is easily distinguishable in both the arterial and the equilibrium phases; as the middle phase is insignificant for our purposes, the most significant images to be processed for feature extraction and the subsequent lesion classification are acquired from the first and last phases.

3.1 Image Segmentation

Preliminarily, we contrasted the acquired images (Fig. 2(a)) to enhance them. By doing this, we rendered all the areas of the CT image clearly distinguishable; the result is shown in Fig. 2(b).

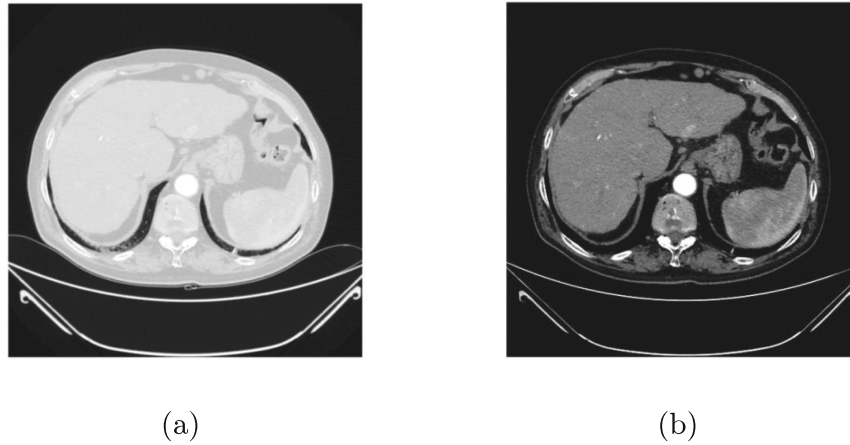


Fig. 2. Image Contrasting: (a) acquired image; (b) result after contrast enhancement

In the second step, images of the arterial phase were processed to isolate the liver from other organs in the same section. The algorithm for isolating the liver is described below.

- Step 1.** The process of liver segmentation starts with selecting the image which contains the maximum extension of the liver (Fig. 3).
- Step 2.** On the starting image (Fig. 4(a)), a morphological opening was applied using a disk ($r = 20$) as structuring element (Fig. 4).
- Step 3.** Starting from image at Step 1 Fig. 5(a), a morphological erosion was applied with the same structuring element. Subsequently, the image is reconstructed using both the starting image and the one coming from the erosion operation. The result is shown in Fig. 5(b).
- Step 4.** On the image from Step 2 Fig. 6(a), a morphological closure was applied using the same structuring element. The result is shown in Fig. 6(b).
- Step 5.** On the image from Step 3, a morphological dilatation was applied using the same structuring element of previous steps; the resulting image was then complemented with the one coming from Step 3. Both images were then used

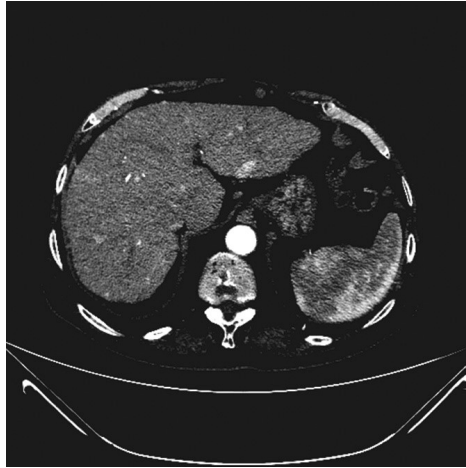
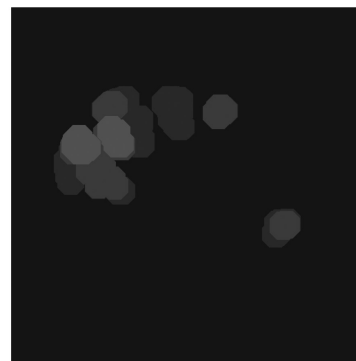


Fig. 3. CT slice with the larger extension of liver



(a)

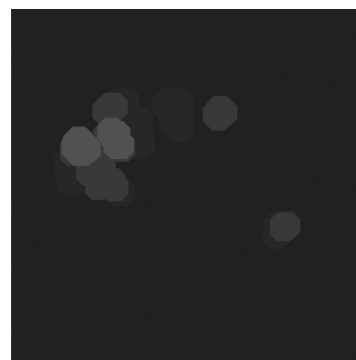


(b)

Fig. 4. Application of morphological opening: (a) starting image; (b) result of morphological opening



(a)



(b)

Fig. 5. Application of morphological erosion and reconstruction: (a) starting image; (b) result of reconstruction



Fig. 6. Application of morphological closure: (a) starting image; (b) result of morphological closure.



Fig. 7. Individuation of regions with high extension: (a) result of image reconstruction after the morphological operations; (b) obtained binary mask

together in the reconstruction. Finally, the resulting image was complemented (Fig. 7(a)) and the connected regions with the larger extension of areas were found on the resulting image (Fig. 7(b)).

Step 6. We applied a morphological closure on the image obtained in the previous step using a 5-by-5 unitary matrix as structuring element; then, a morphological opening was applied using the same structuring element. Subsequently, all the areas with less than 20 pixels were removed. The resulting image is then superimposed on the starting image (Fig. 8).

Step 7. On the image from Step 5, the watershed transformation was applied; this allowed to obtain 2 separated areas containing the liver and the other organ respectively (Fig. 9(a)). The ridge line obtained from the watershed transformation was used to cut off the other organs from the CT image (Fig. 9(b))



Fig. 8. A representation of the starting image with the superimposed areas

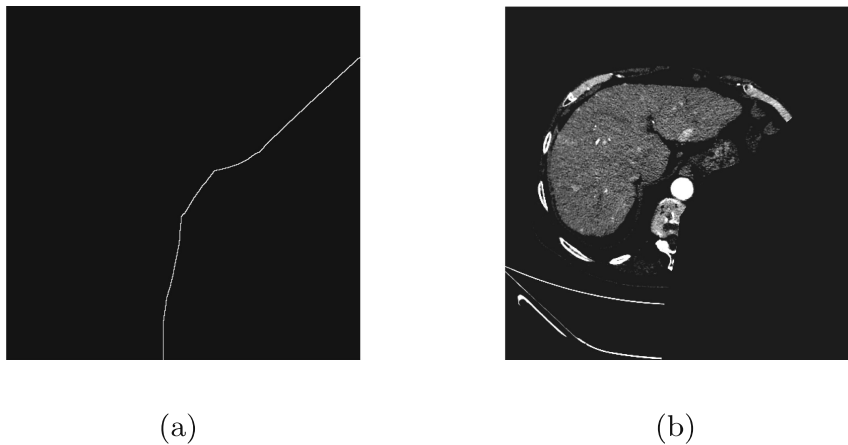


Fig. 9. Image showing the result of: (a) watershed transformation; (b) the application of the obtained ridge line

Step 8. Steps 4 and 5 were repeated on the resulting image; the area with maximum extension was found and the final image which contains the liver was obtained (Fig. 10).

Finally, a mono-seed region growing algorithm [13] was applied for a complete segmentation and isolation of liver; a restrained threshold value enabled us to completely isolate the liver; this, in turn, is clearly separated from the other parts which were not excluded in the previous steps of the algorithm. An example of the final result obtained from the segmentation algorithm could be seen in Fig. 11.

From the segmented liver, a binary mask was obtained which was then applied to all the other slices from the arterial phase of acquisition protocol for obtaining only images that contain parts of liver.

The described processing algorithm, together with the correct execution of the acquisition protocol, results in a correct extraction of the regions of interest (ROIs), including liver lesions (Fig. 12), from the portion of the liver in the CT images acquired during the arterial phase.

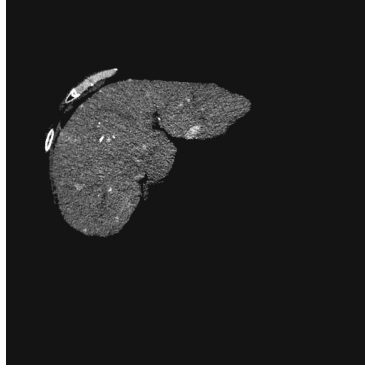


Fig. 10. Image containing the result of segmentation algorithm

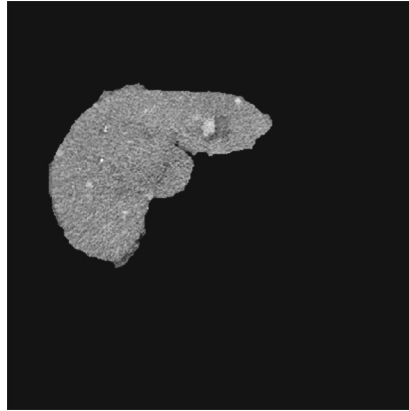


Fig. 11. Completely isolated liver

The segmentation of ROIs containing lesions was realized using a multi-seed region growing algorithm. Actually, the seeds are positioned by a physician using a graphical user interface (GUI) specifically developed for this purpose. After the ROI has been selected, it is automatically reported on the respective slice acquired during the equilibrium phase. By doing this, using some textural descriptors, we could describe the same region but under different conditions during the CT exam.

3.2 Feature Extraction

We used textural features extracted using gray level co-occurrence matrices, as proposed by Haralick [14] and derived from his work [15, 16], in order to describe the ROIs. They were extracted from the ROIs detected in the previous step (in both the arterial and the delayed phases). Specifically, 22 features were extracted from each ROI. Consequently, each entry in the dataset corresponding to a lesion found in 2 slices (acquired in both the arterial and the equilibrium phases) was described by a total of 44 features.

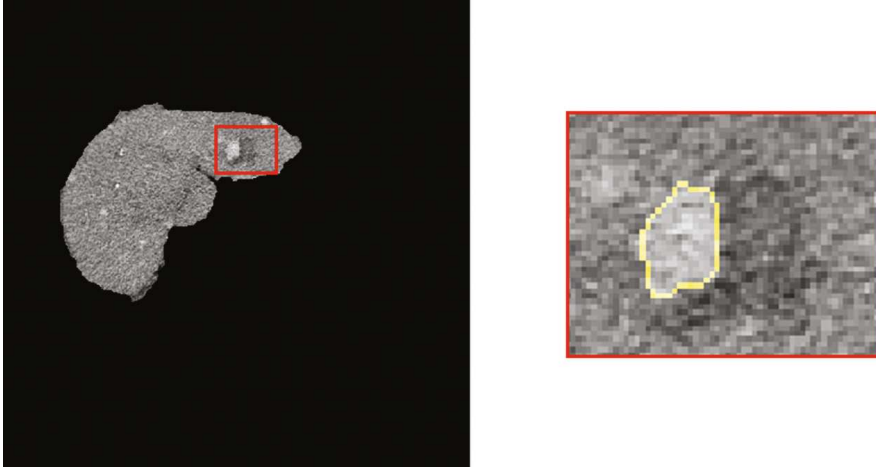


Fig. 12. Lesioned ROI selected on the liver

4 Classification

As stated in Sect. 2.2, we divided patients in two groups, in order to perform a binary classification. The first group contains subjects with well/moderately differentiated nodule; conversely, patients with poorly differentiated nodules are grouped in the second set. In this section, we describe the Artificial Neural Network (ANN) used to classify the degree of lesions between the two groups mentioned above.

4.1 Data Analysis and Preprocessing

Prior to the analysis, we excluded CT images from one patient, because their acquisition was not performed correctly. The resulting dataset was composed of **31** entries with a lesion in the first group (negatives) and **17** samples with a lesion in the second one (positives). Indeed, the performance of machine learning algorithms in medicine is evaluated using predictive accuracy, that is, a measure of diagnostic ability to correctly predict the output. As reported in [17], a small dataset affected by unbalance between classes could undermine the validity of a classification process. Therefore, we balanced the dataset using the Adaptive Synthetic Sampling (ADASYN) algorithm [18], an improved version of the Synthetic Minority Oversampling Technique (SMOTE) [19], in order to increase the number of positive patterns (second group examples) until the amount of negatives ones (first group examples) to improve the learning ability of the classifier. As a result, the final dataset was perfectly balanced, and it shows the same number of samples for each class (**31** positives and **31** negatives), so a total number of **62** samples. Moreover, starting from the initial dataset, the number of features was reduced to:

- 22 features coming from an algorithm of ranking based on the relative entropy, also known as Kullback-Leibler distance or divergence [20];

- 5 features of Haralick, that is, Contrast, Correlation, Energy, Homogeneity and Entropy, which have been used in a previous work for blood vessels and tubules classification [21].

In all the cases mentioned above, the dataset was standardized before classification using the z-score technique [22], in which the aim is to rescale the absolute values in data in an interval centered in 0 and with a variance equal to 1.

4.2 Classifiers

Several approaches for finding the best topology for an ANN classifier with Genetic Algorithms are described in literature; in [23], the authors search the optimal topology of ANN using a multi-objective genetic algorithm (MOGA). In this work, we found the topology of the ANN using a mono-objective Genetic Algorithm (GA) which finds the best topology maximizing the average test accuracy on a certain number of training, validation and test iteration for each topology of the ANN using different permutations of the dataset [24].

For each input dataset, the GA was executed to find the optimal topology in each case. Specifically, the GA finds topology of the ANN in terms of number of hidden layers (ranging from 1 to 3), number of neurons per layer (ranging from 1 to 256 for the first hidden layer, and from 0 to 255 for the other hidden layers), and activation functions in *log-sigmoid* (logsig), the *hyperbolic tangent sigmoid* (tansig), the *pure linear* (purelin) and the *symmetric saturating linear* (satlins), for all the neurons per-single layer.

The performances of the classifiers were evaluated in terms of accuracy (Eq. 1), specificity (Eq. 2) and sensitivity (Eq. 3), using the confusion matrix as follow (Table 1):

Table 1. The configuration of confusion matrix.

Predicted	True	
	Positive	Negative
Positive	TP	FP
Negative	FN	TN

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Specificity = \frac{TN}{FP + TN} \quad (2)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (3)$$

5 Results

The optimal topologies for ANNs specified by the Genetic Algorithm in the three cases are:

Case 1. Balanced dataset with all the 44 features (22 for arterial phase and 22 for equilibrium phase)

ANN with: 4 layers, 117 neurons for the first hidden layer, 199 for the second, 6 for the third and 1 neuron for the output layer. The activation functions found by the GA were *logsig* for all the hidden layers, while the *tansig* function was set for the neuron in the output layer.

Case 2. Balanced dataset with the ranked 22 features

ANN with 3 layers: 16 neurons for the first hidden layer, 14 for the second and 1 neuron for the output layer. The activation functions found by the GA were *logsig* for all the hidden layers, while the *tansig* function was set for the neuron in the output layer.

Case 3. Balanced dataset with the 10 Haralick features (5 for arterial phase and 5 for equilibrium phase)

ANN with: 4 layers, 20 neurons for the first hidden layer, 17 for the second, 13 for the third and 1 neuron for the output layer. The activation functions found by the GA were *tansig* for the first hidden layer, *logsig* for the second, *satlins* for the third, while the *tansig* function was set for the neuron in the output layer.

The ANN training, validation, and test sets were obtained from the input dataset with 60% of samples for training, 20% for validation, and 20% for test set. Specifically, at each iteration, the above sets are obtained through a random permutation of the dataset, maintaining the number of samples of each class balanced. Moreover, the classification thresholds were determined using Receiver Operating Characteristic (ROC) curves [25], by evaluating the True Positive Rate (TPR) against the False Positive Rate (FPR) at various thresholds setting, in order to find the value which could achieve the best discrimination between the two classes.

Results are expressed in terms of mean values, considering 100 iterations, for accuracy, specificity, and sensitivity. In the first case, accuracy was 91.43% (std = 3.04), specificity was 0.974 (std = 0.059), and sensitivity was 0.8543 (std = 0.0201). In the second case accuracy was 92.14% (std = 2.15), specificity was 0.986 (std = 0.043), and sensitivity was 0.857 (std = 0.0001). In the third

Table 2. Confusion matrix for the best case.

Predicted	True	
	Positive	Negative
Positive	TP = 6	FP = 0
Negative	FN = 1	TN = 7

case accuracy was 88.93% (std = 4.80), specificity was 0.95 (std = 0.079), and sensitivity was 0.8286 (std = 0.073).

Due to the small number of samples, the performances in all the cases are quite similar; in particular, the maximum value for the accuracy was the same in the three cases and equal to 92.8571%. The confusion matrix is shown in Table 2.

6 Conclusions

The results of our study show that HCC grading can be discriminated using the extracted features which were discussed in the previous section: the HCC wash-in and wash-out dynamic suggests that this type of lesion could be characterized by analyzing the differences in textural features in both the arterial and the equilibrium phases, ignoring the portal phase. Specifically, the group of feature which enable the best discrimination between the classes consists in the 22 features extracted by the ranking algorithm. Although the number of cases we could analyze is relatively small, the preliminary results obtained in this work and shown in Sect. 5 are indeed very promising. In conclusion, recruiting new patients will result in confirming the results discussed in this paper, and in improving both the method and the outcome of our study.

References

1. Ferenci, P., Fried, M., Labrecque, D., Bruix, J., Sherman, M., Omata, M., Heathcote, J., Piratsivuth, T., Kew, M., Otegbayo, J.A., et al.: Hepatocellular carcinoma (HCC): a global perspective. *J. Clin. Gastroenterol.* **44**(4), 239–245 (2010)
2. Llovet, J.M., Burroughs, A., Bruix, J.: Hepatocellular carcinoma. *The Lancet* **362**(9399), 1907 (2003)
3. Bartolozzi, C., Crocetti, L., Lencioni, R., Cioni, D., Della Pina, C., Campani, D.: Biliary and reticuloendothelial impairment in hepatocarcinogenesis: the diagnostic role of tissue-specific MR contrast media. *Eur. Radiol.* **17**(10), 2519–2530 (2007)
4. Sanyal, A.J., Yoon, S.K., Lencioni, R.: The etiology of hepatocellular carcinoma and consequences for treatment. *Oncologist* **15**(Suppl. 4), 14–22 (2010)
5. Kudo, M.: Multistep human hepatocarcinogenesis: correlation of imaging with pathology. *J. Gastroenterol.* **44**(19), 112–118 (2009)
6. Kudo, M., Chung, H., Haji, S., Osaki, Y., Oka, H., Seki, T., Kasugai, H., Sasaki, Y., Matsunaga, T.: Validation of a new prognostic staging system for hepatocellular carcinoma: the JIS score compared with the CLIP score. *Hepatology* **40**(6), 1396–1405 (2004)
7. Memeo, M., Stabile Ianora, A.A., Scardapane, A., Suppressa, P., Cirulli, A., Sabba, C., Rotondo, A., Angelelli, G.: Hereditary haemorrhagic telangiectasia study of hepatic vascular alterations with multi-detector row helical CT and reconstruction programs. *La Radiologia Medica* **109**(1–2), 125–138 (2004)
8. Stabile Ianora, A.A., Memeo, M., Scardapane, A., Rotondo, A., Angelelli, G.: Oral contrast-enhanced three-dimensional helical-CT cholangiography: clinical applications. *Eur. Radiol.* **13**(4), 867–873 (2003)
9. Choi, J.-Y., Lee, J.-M., Sirlin, C.B.: CT and MR imaging diagnosis and staging of hepatocellular carcinoma: part II. Extracellular agents, hepatobiliary agents, and ancillary imaging features. *Radiology* **273**(1), 30–50 (2014)

10. Bruix, J., Sherman, M.: Management of hepatocellular carcinoma. *Hepatology* **42**(5), 1208–1236 (2005)
11. Kleihues, P., Sobin, L.H.: World health organization classification of tumors. *Cancer* **88**(12), 2887 (2000)
12. Edmondson, H.A., Steiner, P.E.: Primary carcinoma of the liver. A study of 100 cases among 48,900 necropsies. *Cancer* **7**(3), 462–503 (1954)
13. Adams, R., Bischof, L.: Seeded region growing. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(6), 641–647 (1994)
14. Haralick, R.M., Shanmugam, K., et al.: Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **6**, 610–621 (1973)
15. Soh, L.-K., Tsatsoulis, C.: Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Trans. Geosci. Remote Sens.* **37**(2), 780–795 (1999)
16. Clausi, D.A.: An analysis of co-occurrence texture statistics as a function of grey level quantization. *Can. J. Remote Sens.* **28**(1), 45–62 (2002)
17. Mazurowski, M.A., Habas, P.A., Zurada, J.M., Lo, J.Y., Baker, J.A., Tourassi, G.D.: Training neural network classifiers for medical decision making: the effects of imbalanced datasets on classification performance. *Neural Netw.* **21**(2), 427–436 (2008)
18. He, H., Bai, Y., Garcia, E.A., Li, S.: Adasyn: adaptive synthetic sampling approach for imbalanced learning. In: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), pp. 1322–1328. IEEE (2008)
19. Chawla, N.V., Bowyer, K.W., Hall, L.O., Philip Kegelmeyer, W.: SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002)
20. Kullback, S., Leibler, R.A.: On information and sufficiency. *Ann. Math. Stat.* **22**(1), 79–86 (1951)
21. Bevilacqua, V., Pietroleonardo, N., Triggiani, V., Gesualdo, L., Palma, A.M., Rossini, M., Dalfino, G., Mastrofilippo, N.: Neural network classification of blood vessels and tubules based on Haralick features evaluated in histological images of kidney biopsy. In: Huang, D.-S., Han, K. (eds.) *ICIC 2015. LNCS (LNAI)*, vol. 9227, pp. 759–765. Springer, Cham (2015). doi:[10.1007/978-3-319-22053-6_81](https://doi.org/10.1007/978-3-319-22053-6_81)
22. Zill, D., Wright, W.S., Cullen, M.R.: *Advanced Engineering Mathematics*. Jones & Bartlett Learning, Burlington (2011)
23. Bevilacqua, V., Mastronardi, G., Menolascina, F., Pannarale, P., Pedone, A.: A novel multi-objective genetic algorithm approach to artificial neural network topology optimisation: the breast cancer classification problem. In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, pp. 1958–1965. IEEE (2006)
24. Bevilacqua, V., Brunetti, A., Triggiani, M., Magaletti, D., Telegrafo, M., Moschetta, M.: An optimized feed-forward artificial neural network topology to support radiologists in breast lesions classification. In: *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion*, pp. 1385–1392. ACM (2016)
25. Bradley, A.P.: The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recogn.* **30**(7), 1145–1159 (1997)