A SURVEY ON CLASSIFICATION OF LIVER TUMOUR FROM ABDOMINAL COMPUTED TOMOGRAPHY USING MACHINE LEARNING TECHNIQUES

Gunasundari S¹
¹Associate Prof, Velammal Engineering College, Chennai

Swetha R²
²Student, Velammal Engineering College, Chennai

Article DOI: https://doi.org/10.36713/epra6570

ABSTRACT
Pattern recognition is a significant area of research in medicine because many applications like diagnostic systems benefit from it. The aim of this research is to analyze developments of liver cancer detection using machine learning techniques for liver disease. The study highlights how liver cancer diagnosis is assisted using machine learning with supervised, unsupervised and deep learning techniques. Several state of art techniques are compared based on performance measures such as accuracy, sensitivity, specificity. Finally, challenges are also highlighted for possible future work.

KEYWORDS: Machine Learning, Liver, Liver disease, Computer Aided Diagnosis system, Liver Cancer, Computed Tomography

1 INTRODUCTION
Liver disease is any trouble of liver function that causes sickness. There are many different types of liver disease. The benign tumors are not cancerous, and doctors can remove them. The benign tumors cells do not spread to tissues around them or to other parts of the body. There are several types of benign liver tumors; they are hemangioma, hepatic adenoma, cysts, lipoma, fibroma, leiomyoma, and focal nodular hyperplasia (FNH). The most common type of benign tumor is hemangioma. This type of benign tumor consists of masses of twisted congested blood vessels and they start in blood vessels. The second most common tumor of liver is FNH. This tumor is the result of congenital arteriovenous malformation hepatocyte response. Although FNH tumors are benign, it is very difficult to differentiate from liver cancer and doctors remove the tumor sometimes when the diagnosis is unclear.

Primary liver cancer is the sixth most frequent cancer globally. Liver cancer is the second leading cause of cancer death. Cancer cells can invade and damage nearby tissues and organs. Most frequent primary liver cancers begin in hepatocytes (liver cells). This type of cancer is called hepatocellular carcinoma (HCC) or hepatoma. Hepatocellular Carcinoma (HCC) is the most common kind of liver cancer, which causes up to 80% of victims (Stewart et al 2014) Some of the other malignant tumors include cholangiocarcinoma, sarcoma, and hepatoblastoma. According to GLOBOCAN report (Bray et al 2018) that liver cancer is the 6th and 7th causes of demises in men and females respectively.

Advances in medical imaging and image processing techniques have greatly enhanced the interpretation of medical images. Radiographic imaging modalities like Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) help in liver
cancer diagnosis. Medical imaging based on CT is suitable for the detection of liver, kidney, and lung cancer diseases. In fact, a large number of Computer Aided Diagnostic System (CAD) systems have been employed for assisting physicians in the early detection of liver, kidney, breast, and lung cancers.

2 LITERATURE SURVEY

Different types of liver cancer classification using machine assistance have opened up a new research area for early detection of cancer. Many authors have surveyed techniques for liver disease diagnosis. (Fujita et al, 2010; Rathore et al, 2011; Gunasundari et al, 2013; Anwar et al, 2018; Brunetti et al, 2019; Saba, 2020; Nayantara et al, 2020). Chen et al (1998) developed system based on fractal geometry and modified probabilistic neural network for the CT liver image classification. Gletsos et al (2001) proposed a computer-aided diagnostic system for the classification of hepatic lesions from CT images. The use of texture features, the application of dimensionality reduction with sequential forward floating selection (SFFS), and the choice of a classifier consisting of three sequentially placed neural networks, have resulted in a total classification performance of 98%, for four classes of hepatic tissue.

Mougiakakou et al (2003) proposed a CAD system based on texture features and a multiple classification scheme for the characterization of four types of hepatic tissue from CT images has been presented. In Gletsos et al (2003), the classifier module consists of three sequentially placed feed-forward neural networks (NNs). The first NN classifies into normal or pathological liver regions. The pathological liver regions are characterized by the second NN as cyst or “other disease.” The third NN classifies “other disease” into hemangioma or hepatocellular carcinoma. Three feature selection techniques have been applied to each individual NN: the sequential forward selection, the sequential floating forward selection, and a genetic algorithm for feature selection. Result shows that genetic algorithm yields lower dimension feature vectors and improved classification performance.

In Valavanis et al (2004), an ensemble of classifiers has been constructed by using neural network to discriminate four hepatic tissue types: normal liver, hepatic cyst, hemangioma and hepatocellular carcinoma from non-enhanced CT images. Mala et al (2005) extracted Wavelet based texture features and used it to train the Probability Neural Network (PNN) to classify the liver as fatty liver or cirrhotic liver. Mala et al (2006,2008) extracted Biorthogonal wavelet based texture features and used it to train the PNN to classify the liver tumor as hepatocellular carcinoma, cholangiocarcinoma, hepatocellular adenoma and hemangioma.

A classifier based on the support vector machine (SVM) is proposed for automatic classification in liver disease (Huang et al, 2006). Lee et al (2007 a), Lee et al (2007 b) introduced a kernel-based classifier for liver disease distinction of computer tomography (CT) images. In Stavroula et al (2007), for each CT liver ROI, five types of texture feature sets, based on first order statistics, spatial gray level dependence matrices, gray level difference matrices, Laws’ texture energy measures, and fractal dimension measurements, are extracted resulting in a total of 89 features. A Genetic Algorithm based feature selection method was applied to feature sets, when dimensionality reduction was desired, while two alternative Ensemble of Classifiers were investigated consisting solely of NN classifiers, or of a combination of NN and statistical classifiers. A plurality and a weighted voting scheme are used to combine the outputs of the primary classifiers of the examined CAD systems.

Bharathi et al (2008) observed that the Zernike moment features are to be efficient in distinguishing the soft liver tissues into normal and abnormal. In Wang et al, 2009, an experienced radiologist identified ROI with CT liver images. Texture features based on First Order Statistics (FOS), SGLDM, Gray Level Run Length Method (GLRLM), and Gray Level Difference Method (GLDM) are extracted for each ROI. Multi class SVM is used to classify liver diseases into primary hepatic carcinoma, hemangioma, and normal liver.

Kumar et al (2010) proposed a novel feature extraction scheme based on multiresolution fast discrete curvelet transform (FDCT) and employed ANN classifier using FDCT features for computer-aided diagnosis of liver disease. Kondo et al (2011) proposed the hybrid Group Method of Data Handling type neural network algorithm using the artificial intelligence for the medical image diagnosis of liver cancer. Stoean et al. 2011 developed automatic tool to differentiate five degrees of liver fibrosis. Hill climbing algorithm is used for selecting best features. Cooperative coevolutionary algorithm is used to evolve set of rules that is based on training samples and then rules are used for characterizing between different degrees of fibrosis for the test case.

Kumar et al (2012) concluded that contourlet based features performed better than the Gray Level textural features. A tandem feature selection mechanism and evolutionary trained neural network model is proposed for the classification of stages for liver fibrosis within chronic hepatitis C (Gorunescu, 2012). Kumar et al. (2013) proposed liver CAD system to classify HCC and hemangioma. Wavelet coefficients and Contourlet coefficients are reduced using Principal Component Analysis (PCA). The reduced features are fed into PNN for classification of benign tumor from malignant and the accuracy got using CCCM feature is 96.7, which is higher than other features.

In Zhou 2014, data visualization and classification method is presented to distinguish between healthy and diseased livers for early-stage diagnosis of liver.
disease. Glowworm swarm optimization is used for optimizing parameters width and penalty constant of Support Vector Data Description that is used to improve diagnostic accuracy. Gunasundari et al (2016) applied Velocity bounded Boolean PSO for feature selection problem to select elite features from liver cancer data. When applied in liver CAD system, proposed algorithm selects elite features, which is used to classify hepatoma and cholangiocarcinoma as malignant and hemangioma and FNH as benign.

Alahmer et al (2016) proposed CAD system which divides a segmented lesion into three areas, i.e. inside, outside and border areas. Features extracted from the three areas are used to build a new feature vector to feed a classifier for better differentiation between benign and malignant lesions. Sayed et al (2016) proposed a fully automatic CAD system for liver and lesion segmentation and for liver disease diagnosis. The liver segmentation approach is based on a hybrid approach of fuzzy clustering and grey wolf optimisation, while a fast fuzzy c-means technique is employed for lesion segmentation. Shape and texture features are extracted from lesions and employed in a classification stage using a support vector machine classifier.

Chang et al (2017) proposed a computer-aided diagnosis (CAD) system to diagnose liver cancer using the features of tumors obtained from multiphase CT image. Three kinds of features were obtained for each tumor, including texture, shape, and kinetic curve. The texture was quantified using 3-dimensional (3-D) texture data of the tumor based on the grey level co-occurrence matrix (GLCM). Backward elimination was used to select the best combination of features, and binary logistic regression analysis was used to classify the tumors with leave-one-out cross validation.

Kuo (2018) proposed a computer-aided diagnosis system with texture analysis for liver tumors in CT images. Improved PSO algorithm and SVM are combined for reducing features to less than 20 by the proposed two-stage selection method. Meanwhile, fuzzy logic weight value is adopted to determine the final classification result.

Gunasundari et al (2018) proposed Multiswarm Heterogeneous Binary Particle Swarm Optimization algorithm and tested in the feature selection phase of intelligent liver and kidney cancer diagnostic systems to select elite features from the liver cancer data. MHBPSO is a cooperation algorithm, which includes BPSO and its three variants such as Boolean PSO (BoPSO), Self Adjusted Hierarchical Boolean PSO (SAHBoPSO), and CatfishSelf Adjusted Hierarchical Boolean PSO (CSAHBoPSO). The elite feature sets are extracted using SAHBoPSO, CSAHBoPSO, and MHBPSO1 and they are used to classify the liver disease as benign or malignant with minimum error rate. Results show that MHBPSO1 is superior in selecting elite features.

Gunasundari et al 2019 proposed adjusted BPSO (ABPSO) to improve the performance of feature selection and the convergence speed of BPSO. The algorithm is verified in the feature selection module of liver CAD system. Yamashita et al (2020) shows the feasibility of using Convolutional Neural Network (CNN) to assign LI-RADS categories (LR1–5) to liver observations on multiphase CT and MRI from a relatively small sample of images. Li et al (2019) designed a liver cancer diagnosis system that is based on BP neural network after dimensionality reduction by PCA, which combines multiple features and multi-phase information. In Al-Shabi et al (2019), the performance of SVM, MLP and GRNN has been examined in classifying the liver tissues (affected or unaffected) dataset.

Das et al (2019) proposed a new system called as watershed Gaussian based deep learning (WGDL) technique for effective delineate the cancer lesion in CT images of the liver. The segmented textural features were fed to deep neural network (DNN) classifier for automated classification of three types of liver cancer i.e. hemangioma (HEM), hepatocellular carcinoma (HCC) and metastatic carcinoma (MET). Sadeque et al (2019) presents an automated method of detecting liver cancer in abdominal CT images and classifying them using the histogram of oriented gradient - support vector machine (HOG-SVM) algorithm. In Rajathi et al (2019), to classify Chronic liver disease (CLD), the hybrid whale optimization algorithm with simulated annealing (WOA-SA) is used in selecting an optimal set of features. A hybrid ensemble classifier with support vector machine, k-Nearest Neighbor (k-NN), and random forest (RF) classifiers are used to classify liver diseases.

Dong et al (2020) presented the Hybridized fully convolutional neural network (HFCNN) method for liver cancer and lesion identification and segmentation. In Li et al (2020), Liver and tumor segmentation are implemented by a fully convolutional networks using a weighted loss function. Then the output of FCN is combined with the original CT image to construct a 4- channel image data as the input of a 9-layers CNN classifier.

Shapira et al (2020) implemented a weakly-supervised convolutional neural network (CNN) that learns liver lesion localisation without pixel-level ground truth annotations. Regions-of-interest are selected automatically based on the localisation results and are used to train a second CNN for liver lesion classification (healthy, cyst, hypodense metastasis). Lee et al (2020) designed multiple prediction models for 5-year metachronous liver metastasis (5YLM) using combinations of clinical variables (age, sex, T stage, N stage) and top principal components (PCs), with logistic regression classification.

Literature review reveals [Table 1] that only some diseases like ICC and hemangioma are considered by many authors. Diseases like cholangiocarcinoma,
FNH, Fatty liver, cirrhosis and other diseases are not focused by many authors. Therefore, there is a requirement for a liver CAD system to characterize the benign tumor from the malignant tumor, which considers the rare diseases.

Table 1: Performance of existing liver CAD system

<table>
<thead>
<tr>
<th>Author Name</th>
<th>Textural Features and Classifier</th>
<th>Classifier</th>
<th>Disease</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen 1998</td>
<td>detect-before-extract (DBE) system, fractal feature information and the gray-level co-occurrence matrix</td>
<td>Modified PNN</td>
<td>Hepatoma and Hemangioma</td>
<td>accuracy = 83%</td>
</tr>
<tr>
<td>Gletsos 2001</td>
<td>Angular second moment, contrast, correlation, sum of squares, inverse difference moment, entropy, homogeneity, cluster tendency</td>
<td>Feed-Forward NN</td>
<td>Normal liver, hepatic cysts, hemangioma, and HCC</td>
<td>accuracy = 98%</td>
</tr>
<tr>
<td>Gletsos 2003</td>
<td>48 texture characteristics from spatial gray-level co-occurrence matrices- Angular Second Moment, Contrast, correlation, sum of Square – Variance, Inverse Difference Moment, Entropy, Homogeneity, Cluster Tendency</td>
<td>Feed-Forward NN</td>
<td>Normal liver, hepatic cysts, hemangioma, and HCC</td>
<td>accuracy = 97%</td>
</tr>
<tr>
<td>Mougiakako u 2003</td>
<td>FOS, SGLDM, GLDM, TEM, and FDM</td>
<td>4-class NN – Back propagation Algorithm</td>
<td>Normal liver, hepatic cyst, hemangioma, and HCC</td>
<td>accuracy = 97%</td>
</tr>
<tr>
<td>Valavanis 2004</td>
<td>texture features</td>
<td>Two Neural Network (NN) and three statistical classifiers</td>
<td>Normal, hepatic cyst, hemangioma, hepatocellular carcinoma</td>
<td>accuracy = 90.63%</td>
</tr>
<tr>
<td>Mala 2005</td>
<td>Orthogonal wavelet based texture analysis</td>
<td>PNN</td>
<td>Fatty or cirrhosis</td>
<td>accuracy = 95%, sensitivity = 96% and specificity = 94%</td>
</tr>
<tr>
<td>Huang 2006</td>
<td>Auto-Covariance Texture Features</td>
<td>SVM</td>
<td>Malignancy and hemangioma</td>
<td>accuracy 81.7%, sensitivity 75.0% specificity 88.1% positive predictive value 85.7% negative predictive value 78.7%</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Features and Techniques</td>
<td>Classifier/Algorithm</td>
<td>Class</td>
<td>Average AUC of Distinction</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------------</td>
<td>----------------------</td>
<td>------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Lee 2007 a</td>
<td>gray level and co-occurrence matrix features and region-based shape descriptors</td>
<td>SVM</td>
<td>cysts, hepatoma, cavernous hemangioma, and normal tissue</td>
<td>0.91</td>
</tr>
<tr>
<td>Lee 2007 b</td>
<td>gray level co-occurrence matrix, and shape descriptors</td>
<td>kernel-based classifier</td>
<td>cyst, hepatoma and cavernous hemangioma</td>
<td>0.91</td>
</tr>
<tr>
<td>Mougiaakakou 2007</td>
<td>FOS, SGLDM, GLDM, TEM, and FDM Feature Selection : GA</td>
<td>Multilayer Perceptron NN</td>
<td>Normal liver, Hepatic cyst, hemangioma, and HCC</td>
<td>84.96%</td>
</tr>
<tr>
<td>Bharathi 2008</td>
<td>Zernike moments and Legendre moments SFS,SBS</td>
<td>Nearest mean classifier</td>
<td>normal and abnormal with HCC</td>
<td>90.2%</td>
</tr>
<tr>
<td>Mala 2008</td>
<td>wavelet based texture features</td>
<td>PNN</td>
<td>hepatocellular carcinoma, cholangio carcinoma, hemangioma and hepatoadenoma</td>
<td>90.2%</td>
</tr>
<tr>
<td>Wang 2009</td>
<td>FOS, SGLDM, GLRLM and GLDM</td>
<td>Multi Class SVM</td>
<td>HCC, Hemangioma, Normal</td>
<td>97.78%</td>
</tr>
<tr>
<td>Kumar 2010</td>
<td>FDCT</td>
<td>Feed Forward NN</td>
<td>HCC and hemangioma</td>
<td>Curvelet - 93.3% Wavelet – 88.9%</td>
</tr>
<tr>
<td>Stoean 2011</td>
<td>24 indicators like Stiffness, Sex, BMI, Glycemia etc.</td>
<td>CCEA</td>
<td>Different degrees of fibrosis</td>
<td>accuracy = 62.11%</td>
</tr>
<tr>
<td>Gorunescu 2012</td>
<td>25 indicators like Stiffness, Sex, BMI, Glycemia etc. GA: Feature Selector</td>
<td>MLP</td>
<td>Different degrees of fibrosis</td>
<td>accuracy = 65.21%</td>
</tr>
<tr>
<td>Gunasundari 2012</td>
<td>Co-occurrence matrix and Fast discrete Curvelet transform</td>
<td>BPN, PNN and CFBPN.</td>
<td>hemangioma, and HCC</td>
<td>BPN - 96% PNN - 96% CFBPN -96%</td>
</tr>
<tr>
<td>Kumar 2012</td>
<td>FOS, GLCM,CCFOS, CCCMs Feature Selection : PCA</td>
<td>PNN</td>
<td>HCC and hemangioma</td>
<td>FOS – 79% GLCM -86% CCFOS -93% CCCM – 94%</td>
</tr>
<tr>
<td>Kumar 2013</td>
<td>Gray Level, Co-occurrence features, wavelet co-efficient and contourlet coefficient statistics PCA as feature Selector</td>
<td>PNN</td>
<td>Benign from malignant HCC and hemangioma</td>
<td>CCCM -96.7%</td>
</tr>
<tr>
<td>Zhou 2014</td>
<td>13 features Age, Gender, ALT, AST etc. Glowworm swarm optimization</td>
<td>SVDD Support Vector Data Description</td>
<td>Healthy and Unhealthy</td>
<td>accuracy = 84.28%</td>
</tr>
<tr>
<td>Gun-</td>
<td>Textural Features</td>
<td>PNN, SVM</td>
<td>Hepatoma,</td>
<td>accuracy</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Methodology</td>
<td>Model</td>
<td>Disease Description</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------</td>
<td>----------------------------------------------------------</td>
</tr>
</tbody>
</table>
3 CONCLUSION

In recent years, significant efforts have been made towards the development of the liver CAD system. Machine learning plays a vital role in liver CAD because organs may not be signified precisely by a simple equation or formulae. This study contributed an extensive literature review on the development of liver CAD systems from abdominal CT. The literature review revealed that performance of CNN in recent liver CAD system is good. Only some diseases like HCC, Cyst and hemangioma are considered by many authors. Diseases like cholangiocarcinoma, FNH, Fatty liver, cirrhosis and other diseases are not focused by many authors. Therefore, there is a requirement for a liver CAD system to characterize the benign tumor from the malignant tumor, which considers the rare diseases. In future deep learning will play a major role in liver CAD system. In this future work, we aim to find liver CAD system using deep learning model which characterize benign and malignant tumor from abdominal CT that consider rare liver diseases also.

4 REFERENCES

International Journal of Computer Applications 74(11), 7-12 (2013).


40. Shapira, N., Fokuhl, J., Schulthei, M., Stefanie, B., Kopp, FK., Pfeiffer, D., Dangelmaier, J., Pahn, G.


