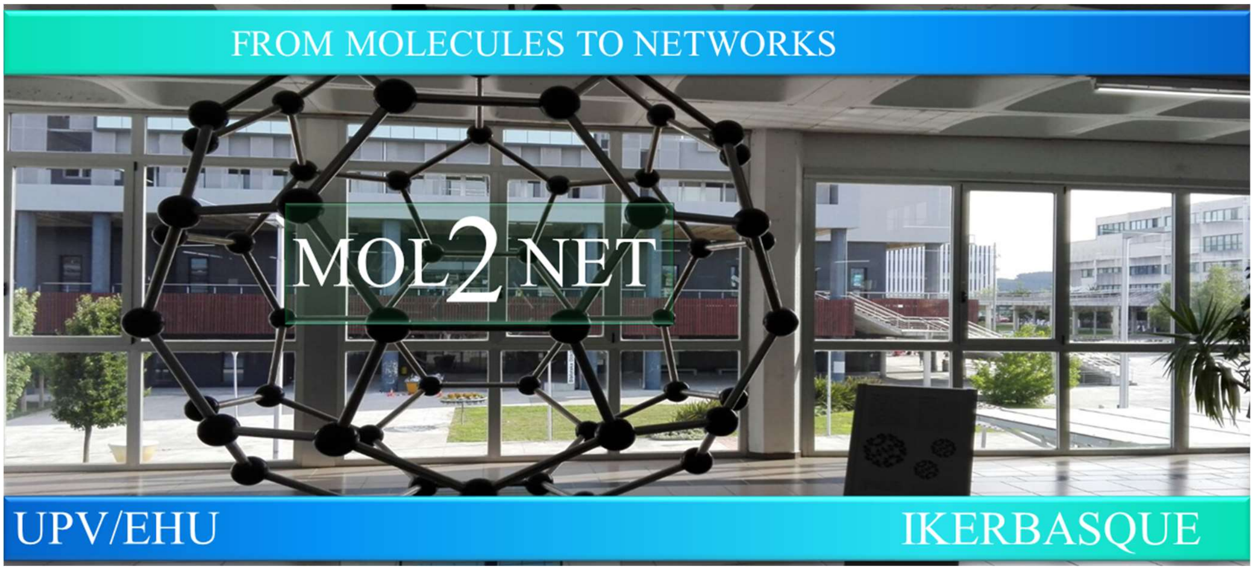
MOL2NET'22, Conference on Molecular, Biomedical & Computational Sciences and Engineering, 8th ed.



Machine Learning Methods Using Texture Feature Selection in Diagnosis of Liver Cancer

Samreen Naeem [*, a], *Aqib Ali* [a], *Sania Anam* [b], *and Muhammad Zubair* [a]

[a] *College of Automation, Southeast University, Nanjing, China.*
[b] *Govt Associate College for Women Ahmadpur East, Bahawalpur, Pakistan.*

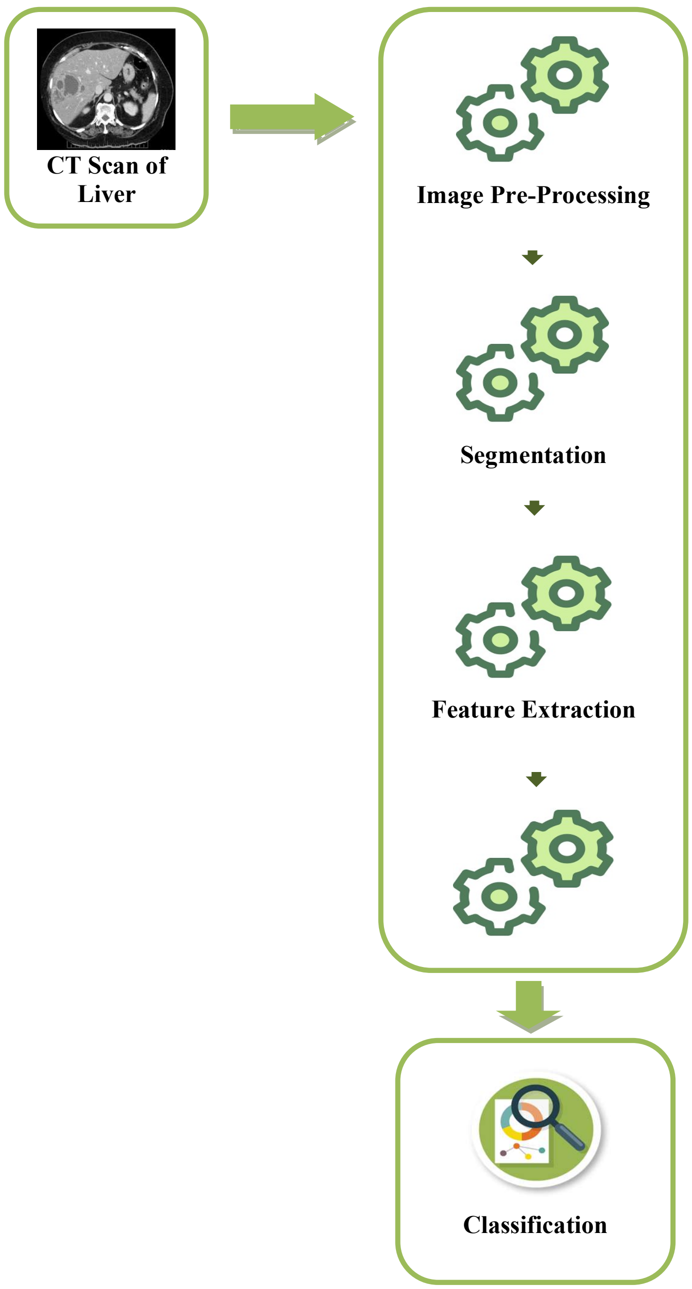
. * *Corresponding author: samreencsit@gmail.com*

Abstract.

The liver is essential to a wide variety of physiological and metabolic activities that take place in our bodies. A lack of liver function or liver dysfunction can lead to various health issues, the most serious of which is the illness. Early identification of liver illness can help shorten the treatment duration and minimize liver damage by decreasing the usage of drugs that are not essential. The application of machine learning techniques in medicine has shown promising results in illness diagnosis due to advances in technology. This research aimed to discover the essential characteristics of the data set, using textural feature selection methods so that they could be applied in the early diagnosis of liver disorders. This was done in order to diagnose liver failure disease. The model in machine learning methods has been improved, which has led to an improvement in the success rate of illness detection. The results obtained were compared with the findings of other research published in the literature that used the same data set. The diagnostic success rate for liver failure illness, as determined by applying machine learning algorithms known as Decision Trees (DT), was 97.16%. It is hoped that the developed models may assist medical professionals in the early detection of liver illness.

Keywords: Liver Cancer, Texture Features, Machine Learning, Decision Tree.

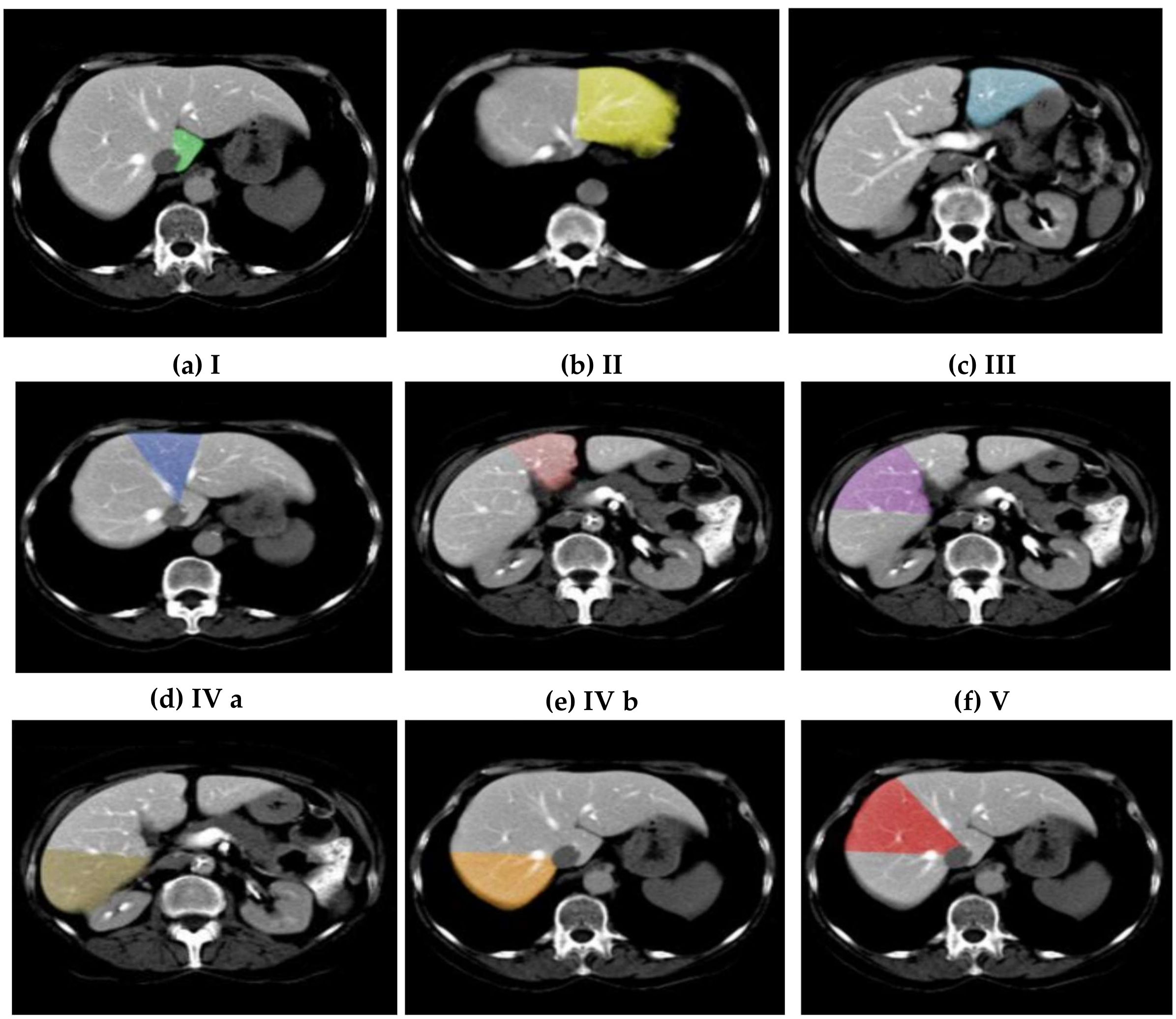
Graphical Abstract



Introduction

The liver is the human body's most significant and biggest internal organ. It is responsible for many fundamental bodily processes, including the elimination of toxins and hormones, the production of proteins, and the filtering of blood. The illness that poses the most significant risk is hepatocellular carcinoma, often known as liver cancer. Hepatocellular carcinoma (also known as HCC) is the most frequent hepatic cancer, and its proportion accounts for 80 percent of all hepatic malignancies. Alcohol, tobacco use, obesity, and other risk factors are some of the many things that may lead to liver cancer. In its early stages, liver cancer is notoriously difficult to diagnose [1]. The diagnostic process in medicine is very significant and must be carried out correctly and quickly. When investigating extreme conditions, there is always the possibility of making a mistake. As a result, an intelligent system would be developed to lower the likelihood of making a mistake in diagnosis [2].

Human liver can be divided into eight functional units. This division is due to the structure of vessels. There are number of approaches for liver segmentation. But in Europe, mostly used approach is Couinaud's segmentation. Each segment is functionally independent and each effected segment can be treated separately. Due to this segmental structure, liver is regeneratable. Liver is comprised of two lobes: The Left Lobe and The Right Lobe. Each lobe consists of several small segment which are known as lobules [3].



(a) I **(b) II** **(c) III**

(d) IV a **(e) IV b** **(f) V**

(g) VI **(h) VII** **(i) VIII**

Figure 1: Imagining of different liver segments.

The left lobe is smaller than the right lobe. Left liver vein divides the left lobe into medial and lateral sectors. The lateral sector is composed of segments II and III. The medial sector makes the segment IV, which is further divided into two segments: the upper segment (IVa) and lower (IVb) segment. The right lobe is divided by the right haptic vein into posterolateral and paramedian sections [4]. These parts are more divided into lower and upper segments by the transversal plane. The right lobe is divided into four segments: VIII, VII, VI and V that are indexed anticlockwise. The segment I is an exceptional segment which is called lobus caudatus, which has its dedicated vascularization, Show in Figure 1.

The field of computers, known as image processing, is gaining more and more relevance in modern times. It is helpful in various fields and may be used in the business sector to control production, in safety systems to evaluate people's biometrics, and in other contexts. Diagnostics, patient monitoring, and the planning of treatments are just some of the medical specialties that benefit significantly from computer-assisted analysis of medical images [5]. The growing number of medical images that will need to be processed in clinical settings is primarily attributable to the fact that the population is becoming older and that current imaging technologies are becoming more widespread. There is a significant unmet need for software solutions that expedite the process of analyzing a medical image and turning it into something valuable and productive. An essential component of this sector is the medical image, which uses various image-processing methods and forms of artificial intelligence to solve issues. This research presents a computerized diagnostic system that adapts to medical imaging to locate and identify malignancies in the liver [6].

Machine learning teaches a computer to do a specific job (in this case, image processing) by presenting it with a data set designed specifically for that purpose. Machine learning uses several different algorithms and methods. We can analyze any of these methods to see which would provide the most effective results for image categorization [7-8]. The results also rely on the kind of image processing you are prepared to adopt since each type has intrinsic qualities. For instance, there is a good chance that the cross-entropy loss function will produce better results in image processing than other loss functions [9].

Materials and Methods

This research aims to propose an intelligent system that classifies liver lesions into two classes using CT images. An image dataset that has two categories of liver lesions was collected. These two categories are i) Benign liver lesions and ii) Malignant liver lesions. The proposed framework defines the complete steps of liver lesion classification and fusion of CT data. This process is defined into six basic steps.

- The first step is image data collection. In this step, CT images were collected from the Radiology Department of Bahawal Victoria Hospital, Bahawalpur, Pakistan.
- The second step is data preprocessing. This involves cleaning and noise removal. That will help to standardize image data of liver lesion varieties. We will examine the qualitative nature of data and extract useful information to help the present study.
- The third step is segmentation. In this step, the liver lesion is separated from the normal liver in the image. Then regions of interest are taken from the segmented areas.
- The fourth step is feature extraction. In this step, different texture features are extracted, like "Co-Occurrence Matrix Features," "Histogram features," "Wavelet Features," and "Run Length Matrix features," using an image processing tool named "Mazda."
- The fifth step is feature optimization. In this step, only those features are selected that will have more significance with higher accuracy.
- After all, the above steps, the dataset is prepared. Now, the sixth step is to apply Naïve Bayes (NB) classifier on the dataset to get the classification results using WEKA Software.

Results and Discussion

- Classification model: Decision Trees (DT)
- Time taken to build the model: 0.34 seconds
- Test mode: 10-fold cross-validation

Table 1: DT Classifier Summary

Total Number of Instances	600	
Correctly Classified Instances	583	**97.1667 %**
Incorrectly Classified Instances	17	**2.8333 %**
Kappa statistic	0.9433	
Mean absolute error	0.0313	
Root mean squared error	0.1404	
Relative absolute error	6.2585%	
Root relative squared error	28.0711 %	

Table 2: DT Classifier Detailed Accuracy

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	Class
0.953	0.01	0.99	0.953	0.971	0.944	0.998	**Normal**
0.99	0.047	0.955	0.99	0.972	0.944	0.998	**Abnormal**
0.972	0.028	0.972	0.972	0.972	0.944	0.998	**Weighted Avg.**

Table 3: Confusion Matrix using DT Classifier

Classified as	A	B
A = Normal	286	14
B = Abnormal	3	297

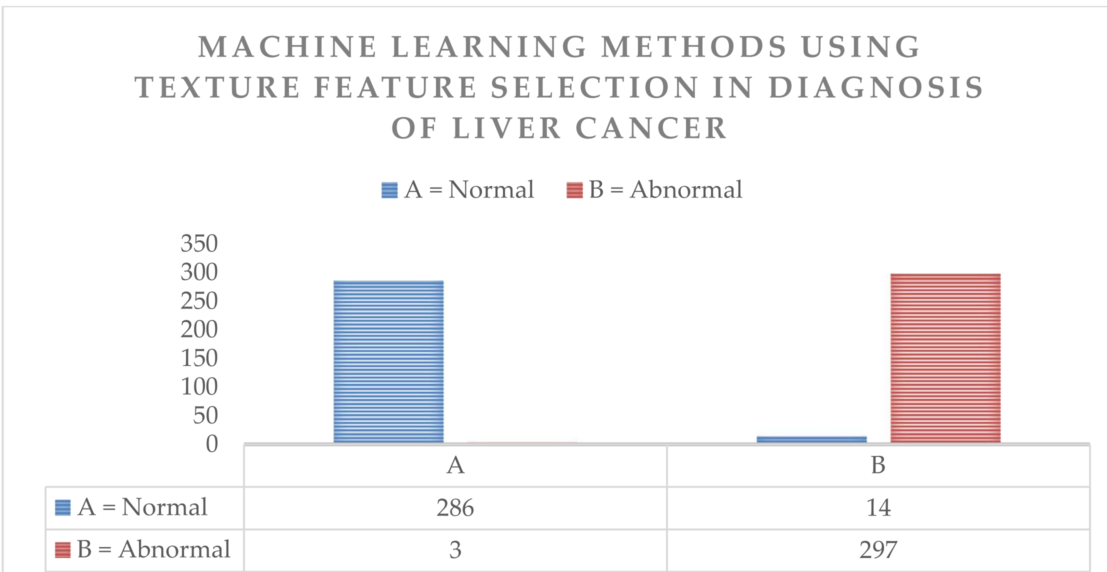


Figure 2: Accuracy of Dataset using DT Classifier (10-Fold)

Conclusions

During this line of research, an intelligent system was provided for categorizing liver lesion types. This system used CT scans, texture analysis, and machine learning (ML) classifiers to determine the various kinds of liver lesions. The goal of this research was to discover and pick relevant optimum features and identify suitable classifiers so that a new classification system may be developed. In addition to this, one of the goals was to locate appropriate classification systems. The results of this analysis lead the investigators to conclude that the use of Decision Trees (DT), which offers an accuracy of 97.16%, can yield more accurate and precise outcomes than what was found before. After all of the examinations were finished, a comparison of CT scans was carried out.

References

[1]. Naeem, S., Ali, A., Qadri, S., Khan Mashwani, W., Tairan, N., Shah, H., ... & Anam, S. (2020). Machine-learning based hybrid-feature analysis for liver cancer classification using fused (MR and CT) images. Applied Sciences, 10(9), 3134.

[2]. Chen, M., Zhang, B., Topatana, W., Cao, J., Zhu, H., Juengpanich, S., ... & Cai, X. (2020). Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. NPJ precision oncology, 4(1), 1-7.

[3]. Trivizakis, E., Manikis, G. C., Nikiforaki, K., Drevelegas, K., Constantinides, M., Drevelegas, A., & Marias, K. (2018). Extending 2-D convolutional neural networks to 3-D for advancing deep learning cancer classification with application to MRI liver tumor differentiation. IEEE journal of biomedical and health informatics, 23(3), 923-930.

[4]. Kiani, A., Uyumazturk, B., Rajpurkar, P., Wang, A., Gao, R., Jones, E., ... & Shen, J. (2020). Impact of a deep learning assistant on the histopathologic classification of liver cancer. NPJ digital medicine, 3(1), 1-8.

[5]. Ali, A., Mashwani, W. K., Naeem, S., Uddin, M. I., Kumam, W., Kumam, P., ... & Chesneau, C. (2021). COVID-19 infected lung computed tomography segmentation and supervised classification approach.

[6]. Ali, A., Qadri, S., Khan Mashwani, W., Kumam, W., Kumam, P., Naeem, S., ... & Sulaiman, M. (2020). Machine learning based automated segmentation and hybrid feature analysis for diabetic retinopathy classification using fundus image. Entropy, 22(5), 567.

[7]. Azer, S. A. (2019). Deep learning with convolutional neural networks for identification of liver masses and hepatocellular carcinoma: A systematic review. World journal of gastrointestinal oncology, 11(12), 1218.

[8]. Das, A., Acharya, U. R., Panda, S. S., & Sabut, S. (2019). Deep learning based liver cancer detection using watershed transform and Gaussian mixture model techniques. Cognitive Systems Research, 54, 165-175.

[9]. Zhang, Z. M., Tan, J. X., Wang, F., Dao, F. Y., Zhang, Z. Y., & Lin, H. (2020). Early diagnosis of hepatocellular carcinoma using machine learning method. Frontiers in bioengineering and biotechnology, 8, 254.