CHRONIC HEPATITIS-C DETECTION INTERPRETATION USING MACHINE LEARNING ALGORITHM AND WITH CLINICAL FINDINGS

[1]K. VIDYA SAGAR, [2]CH.SUMA , [3] SURYA PRASADA RAO BORRA, [4] M.SUMALATHA,
[5]JAMMALAMADUGU RAVINDRANADH, [6]LAKSHMI RAMANI BURRA

[1]Department of EIE, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad
[2]Department of CSE, SRKR Engineering College, Bhimavaram, India
[3]Department of ECE, PVP Siddhartha Institute of Technology, Vijayawada, India
[4]Department of ECE, Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh. India
[5]Department of ECE, RVR & JC College of Engineering, Guntur, Andhra Pradesh, India.
[6]Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India.
.
Email: suryaborra1679@gmail.com

ABSTRACT

Hepatitis - C infection is spreading very rapidly. The infection rate is alarming. The infected liver leads to liver fibrosis and then leads to cirrhosis. The fibrosis and cirrhosis stage are associate the human life. This paper considered CT images for better interpretation and classification. The RBG image is converted to grey scale image. The contrast levels of the image are improved using CLAHE algorithm. The image quality parameters AMBE, PSNR, MSE, MD and MAE are analysed. The contrast adjusted image is processed for banalization using thresholding mechanism then CNN classifier is applied to classify the severity of liver fibrosis, and cirrhosis. The liver profile clinical findings aspartate aminotransferase enzyme (AST) and alanine transaminase (ALT), INR and Albumin total bilirubin are correlated with the machine interpreted results. caudate right lobe ratio of CLD/ RLD is also estimated to interpret the severity of the fibrosis and cirrhosis. The proposed methodology is more reliable for meticulous interpretation of liver inflammation with hepatitis-C virus.

Keywords: *Hep-C Infection, CT Images, CNN Classifier, Liver Fibrosis, Liver Cirrhosis.*

1. INTRODUCTION

With an increase in technology and rapid lifestyle changes, the health issues are significantly alarming. Nowadays, Intensified health issues like hepatitis-C (HEP-C) is increasing across the world and in India too. Hepatitis C is a viral infection that causes inflammation of the liver, sometimes causing severe liver damage.

Hepatitis C virus (HCV) infection affects more than 170 million people worldwide, with 71 million of them having a chronic infection. The Eastern Mediterranean region and Europe have the highest prevalence (2.3 percent and 1.5 percent, respectively), with the rest of the world falling anywhere between 0.5 percent and 1.0 percent. HCV prevalence is reported to be 0.38 percent in Jeddah City, Saudi Arabia. 58 million people have hepatitis C, 75% of hepatitis C patients live in low and middle-income countries, and About 50% of hepatitis C patients live in four countries: China, Pakistan, India, and Egypt. 79% do not know they are infected Only 13% were treated 2.3 million people have both HIV and hepatitis C 15-30% of chronically infected people develop cirrhosis within 20 years 800 people die every day.

During inspiration state and expiration state, the contour extraction is done. From the CT image, the number of voxel elements are counted to predict the top of the liver. The liver portion is extracted from the CT images using the CT value difference between neighbourhood voxel elements. The extracted liver portion shows meticulous morphological coincidence. The shift direction of each voxel element and its direction is considered to determine the deformation of the voxel elements of the infected liver portion from CT images. The

degree of liver fibrosis is classified into three groups. No fibrosis, i.e. normal, bridging fibrosis and bridging fibrosis with distortion and liver cirrhosis [1].

2. METHODOLOGY

Machine learning algorithm CNN augment with Tensor flow and keras is applied on CT images. The image contrast levels are adjusted using CLAHE methodology. With OTSU approach, the threshold value is determined to differentiate the white and black pixels.
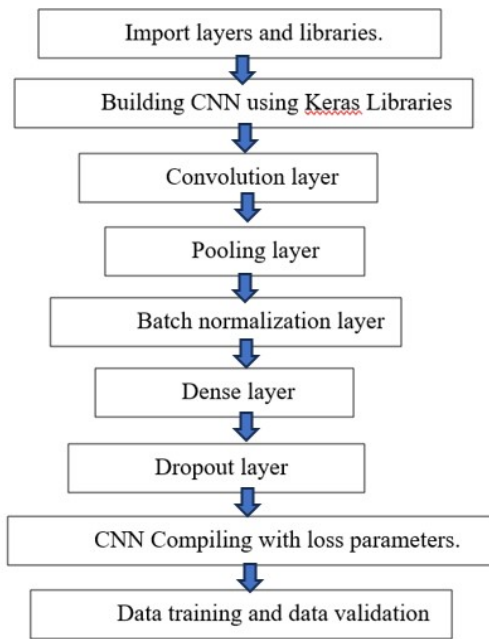

Fig.2.1. Image Classification Procedure Using KNN

During pre-processing stage, the input image is scaled down to minimize the pixels of Hep-C image to improve the maximum number of input nodes. so that the network training time is minimized. Bi-linear interpolation is applied to resize the image.

Image n= cv2. Resize (test_image, (128,128))
code is used for resizing the image.

To equalize pictures, Contrast Limited Adaptive Histogram Equalization (CLAHE) is used. The over-amplification of the contrast is handled with CLAHE. Instead of processing the full image, CLAHE works on discrete sections of the image

called as tiles instead of the full image. Bi-linear interpolation augment with adjacent tails used to curtail the false boarders of HEP-C images. The clip Limit sets the threshold value to '40'for contrast limiting. The image is divided into tiles for applying CLAHE
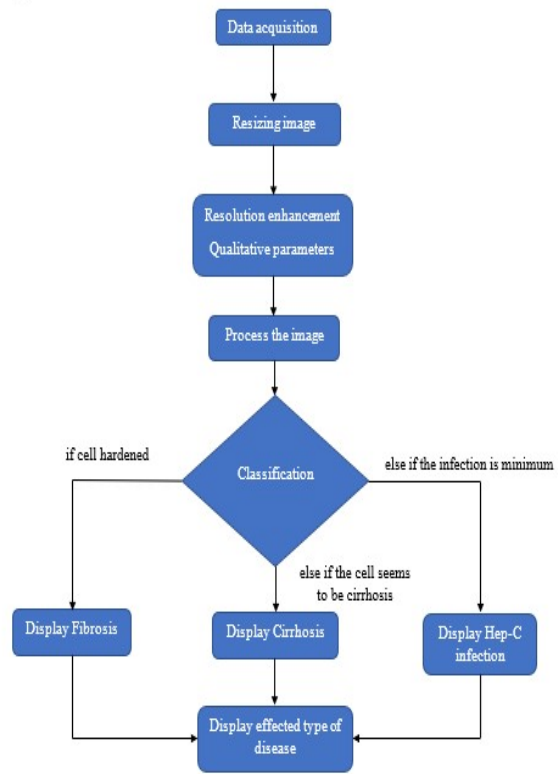

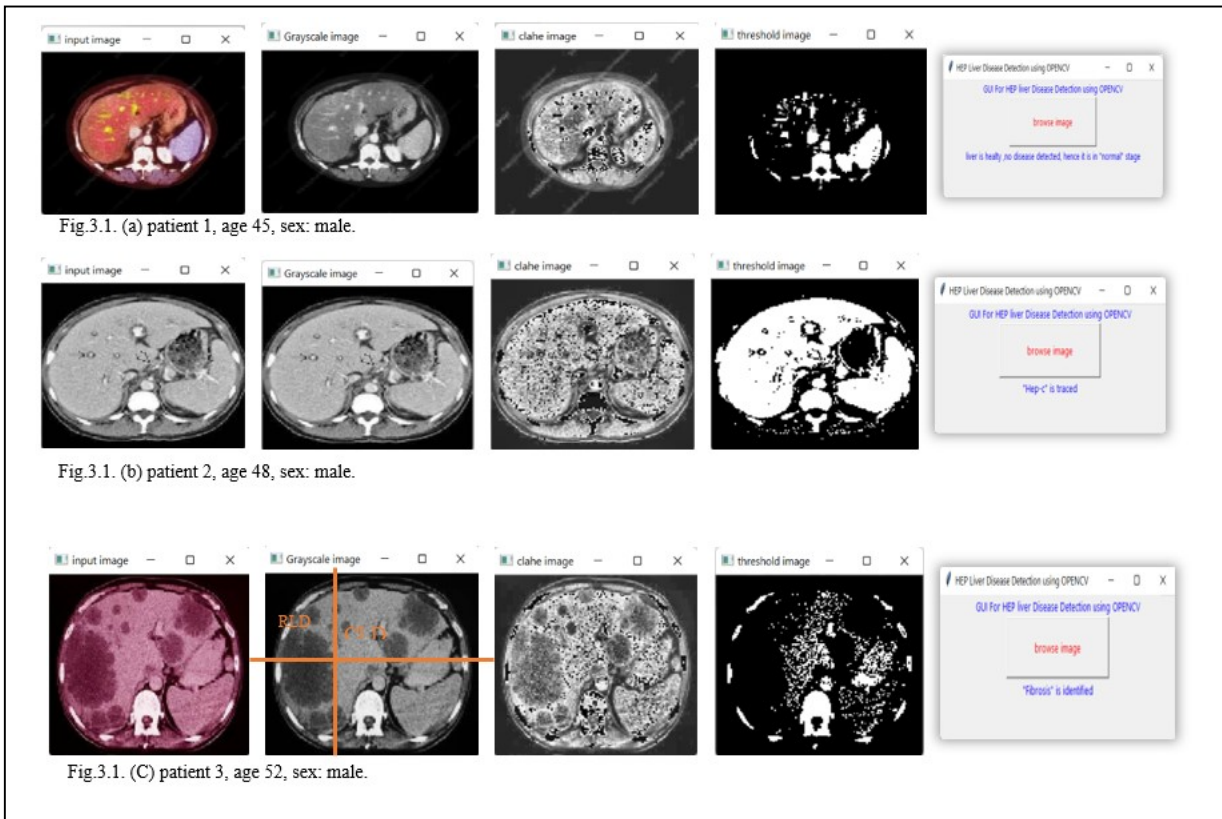*Fig.2.2. The Methodology To Process The Image To Determine The Type Of Disease.*

CNN is used to improve the processing speed and to reduce the images for predicting the HEP-C virus infection without losing original features of an input image. The CNN final layers are fully connected layers. These layers are labelled as dense layers. Keras used to create these layers using Dense () class. Multiple fully connected layers are used to create multilayer perceptron in CNN. Convolution layer is an integral part of CNN used to reduce the high dimensionality of HEP-C CT images for processing, without losing its information[1][2]. The input image is convoluted using convolution operators and the output information is transferred to various channels of the convolution layer. Maximum value is selected from every pool using max pooling layer. The
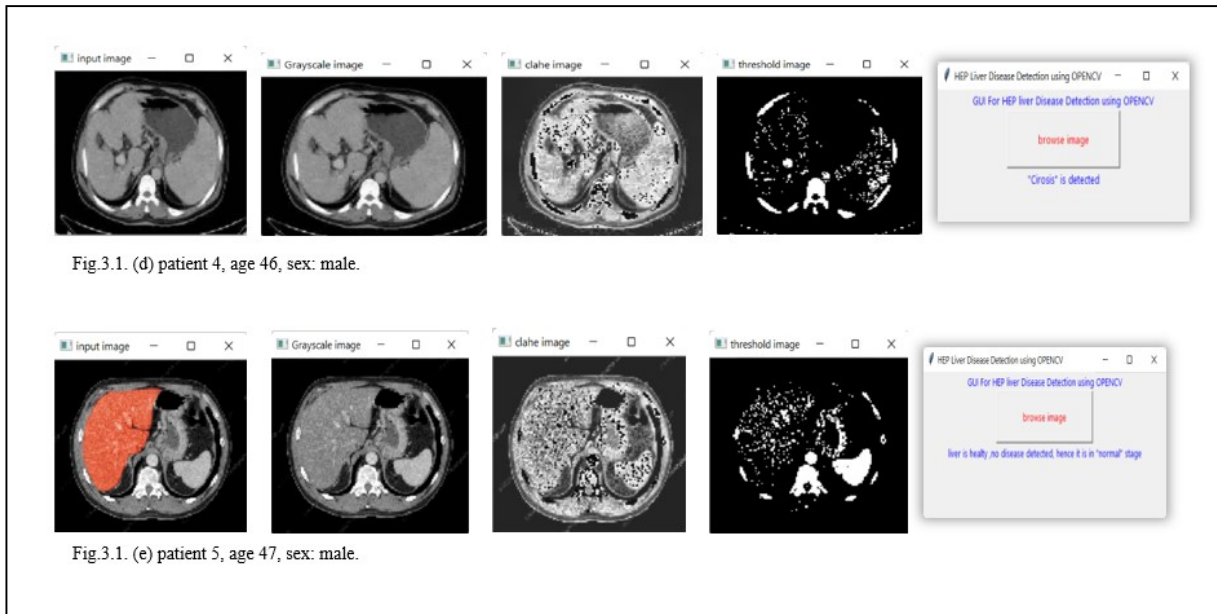
outstanding features of the feature map is retained with max pooling. With max pooling the final image is much sharper than the original image. A continuous Linear vector is obtained by converting all the 2- dimensional arrays from pooled feature maps using flattening. The flattened matrix is applied to fully connected layer to classify the HEP-C images. The resultant output of the convolution layers used to classify the HEP-C images using Dense layer. In keras dense layer all the neurons have tightly coupled each other. Each neuron receives input from all other neurons. Batch norm function is applied to improve the training speed of HEP-C images. Re-scaling and re-cantering are considered for batch normalization. With batch normalization, the accuracy is increased. During training, certain nodes of the network may be disregarding. Dropouts ensure that no node is co-dependent with other node so that the overfitting can be eliminated. Applying dropouts to train HEP-C images restricts to consider the data point more than essential in each dataset.

'Adam' used as an optimizer, categorical cross-entropy as a loss factor, and considered accuracy as a metrics factor. [2][3]

3. RESULTS AND DISCUSSION

The HEP-C CT image is converted into grayscale image and the image is converted into grey scale image at stage-1 and CLAHE algorithm is applied to the grey scale image. Threshold function is applied for meticulous estimation to determine the type of the disease. The clip limit is tuned to 41 to regulate the contrast limiting threshold value. The grid size is tuned to 89 to divide the image into tiles. While improving the contract of CT HEP-C image, both distortion and enhancement are considered in balance position. Over enhancement and significant distortion impacted on contrast enhancement. Absolute mean brightness error (AMBE) is estimated for each image.

Fig.3.1. (a) patient 1, age 45, sex: male.

Fig.3.1. (b) patient 2, age 48, sex: male.

Fig.3.1. (C) patient 3, age 52, sex: male.

Fig.3.1. (d) patient 4, age 46, sex: male.

Fig.3.1. (e) patient 5, age 47, sex: male.

With CLAHE the AMBE values are low, which represents the image contrast is significantly improved. The peak signal to noise ratio (PSNR) is estimated. Significantly good PSNR value is achieved during contrast enhancement of the image represented in table 3.1. Mean square error (MSE) value is estimated for each image. MSE values are less than 1 show the contrast adjustment is significantly good shown in table 3.1. Maximum difference (MD) and mean absolute error (MAE) values are estimated. Table.3.1. shows MSE values are less than 1, which represents the contrast adjustment is good. [4]

Table.3.1. Image Contrast Enhancement Qualitative Parameters Of The Images With CLAHE

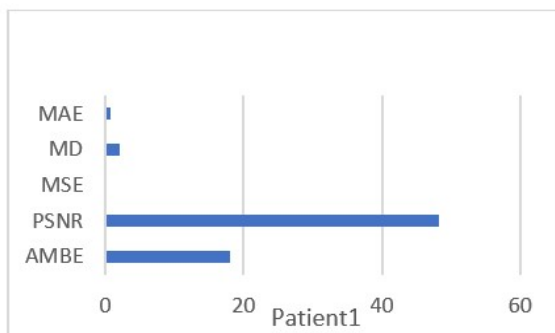| Image	AMBE	PSNR	MSE	MD	MAE
Patient1	18.057	48.23	0.0235	2.0	0.7142
Patient2	21.092	51.32	0.0143	4.2	0.8242
Patient3	17.053	46.02	0.0152	6.7	0.6270
Patient4	19.152	48.93	0.0114	4.5	0.7301
Patient5	21.231	52.44	0.0312	1.9	0.9102

4. EXISTING METHOD
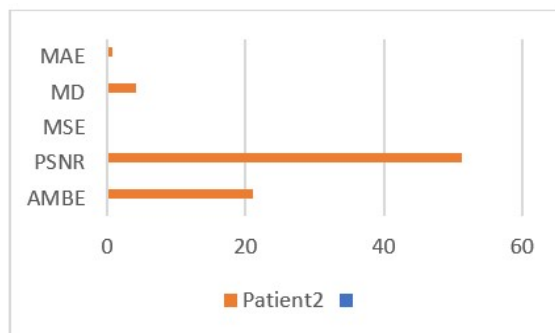

Fig. 3.2 Quality Parameters for Fig. 3.1(a)

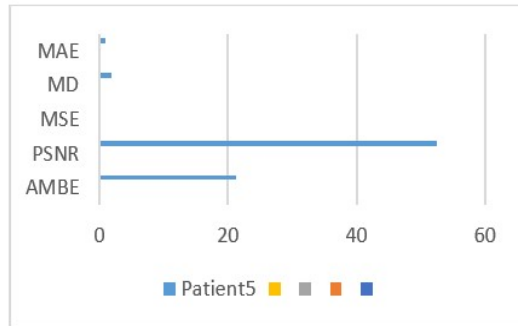Fig. 3.3 Quality Parameters for Fig. 3.1(b)

Fig. 3.6 Quality Parameters for Fig. 3.1(e)

the traces of cirrhosis. The results with machine learning methodology interprets similar findings and are more coincident with the clinical findings.

Alanine transaminase (ALT) values for patient 1 is 54.12, patient 2 is 62.96, patient 3 is 75.05 and for patient 4 is 78.34. The upper reference value of alanine transaminase value is 56. ALT value for patient 1 is 54.12. this value lies within the reference level. But patient 2 the ALT readings are significantly elevated to 12.42 percentage. This represents bridging fibrosis of liver. But patient 3 clinical findings show 34.01 percentage elevated

Table3.2. Liver Profile clinical findings

Image	AST	ALT	TB	Alb	BUN	INR	ALK	Pt	Cr	ALP	APP
Patient1	38.2	54.12	1.18	4.2	20.21	1.12	118.03	2267.6	1.02	118	9.5
Patient2	50.23	62.96	1.83	2.76	18.79	1.24	160.15	2244.0	1.12	120	6.19
Patient3	73.15	75.05	2.63	2.46	26.05	1.38	152.04	1546.3	1.12	152	3.80
Patient4	75.43	78.34	2.89	2.16	29.60	1.58	169.68	1676.7	1.34	167	3.62
Patient5	32.12	45.0	0.89	3.78	16.78	1.02	54.67	2498.9	0.8	65	9.44

The images are taken from ABC radiology department augment with Sateesh gastroenterology. The machine learning results are correlated with liver function test (LFT) results. Aspartate aminotransferase enzyme (AST) and alanine transaminase (ALT) are significant measures of Hep-C viral inflammation. Normal ranges of AST are 5-40 units of serum and ALT are 7 to 56 units of serum., The upper limit values are used to determine the elevated levels of the liver enzymes. The highest levels may be the indication of chronic Hep-C or fibrosis or cirrhosis. AST value for patient 2 is 50.23, patient 3 is 73.15 and patient 4 is 75.43 US. Based on the clinical findings the aspartate aminotransferase (AST) value for patient 2 is moderately high value represents an early stage 2 of hepatitis -C viral infection.[5].

For patients 3 and 4 the AST readings are significantly elevated. Patient 3 value is representative of grade 3 fibrosis. Patient 3 readings are indicative of stage 3 fibrosis and more significantly the elevated results of patient 4 shows

And also, PT, INR and Albumin are better readings to interpret the Stage of hepatitis-C. Albumin normal range is 3.5 g/dL to 5.0 g/dL. Patient 1 value 3.84g/dL is lies within the reference range. This represents the patient 1 liver image is not inflammation with any decease these clinical findings are correlated with ML driven methodology. Patient 2 clinical findings of albumin is 2.76g/dL represents the liver is not able to produce required proteins, this is due to chronic inflammation of the liver. Lowering albumin i.e., 2.76g/dL represents the stage 2 fibrosis. This value is further lowered for patient3, i.e.,2.46g/dL. This value represents stage 3 fibrosis. This value is further lowered to 2.14 for patient 4. This represents the chronic liver inflammation and stage 4 fibrosis or may be cirrhosis. These clinical findings required meticulous correlation with radiology images. The proposed machine learning approach is meticulously detected each radiology CT image on par with the clinical pathology findings. Total bilirubin value is also significant. This value for normal liver is 0.1mg/dL to

1.2mg/dL. Lower reference value of bilirubin is not a significant concern, but elevating levels indicate that malfunction of the liver to clear the total bilirubin. This indication is asymptotic chronic liver disease, or liver may be damaged. For patient 2, patient 3 and patient 4 the total bilirubin values are excessively elevated with the upper reference value. The clinical findings for patient 2 are 1.83. The threshold value for stage 2 fibrosis is 1.35. but the actual value is exceeding the threshold value. The radiology CT image interpreted that the patient 2 is infected with hepatitis-C virus.
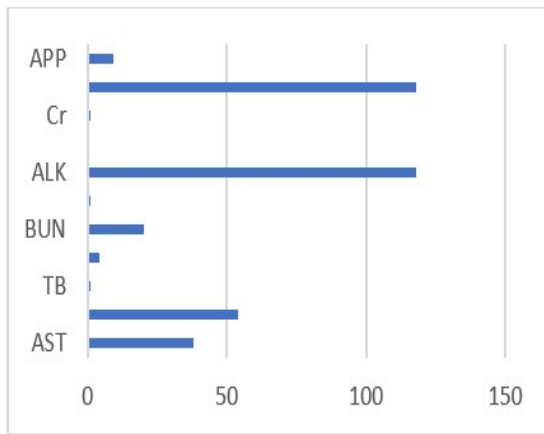

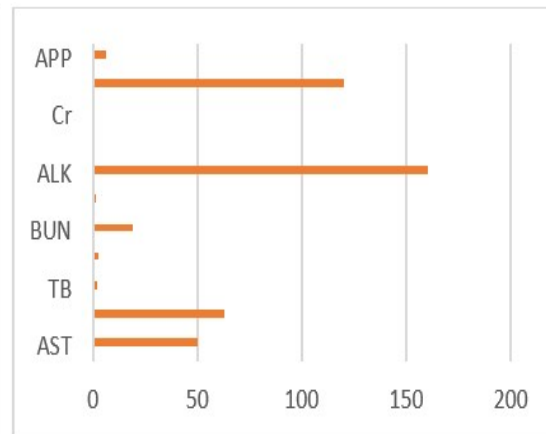Fig. 3.7 Liver Profile findings for Patient 1

Fig. 3.8 Liver Profile findings for Patient 2

But the total bilirubin value shows, the patient is infected with HEP-C and with fibrosis. Patient 3 total bilirubin value is 2.63 and for patient 4 the value is 2.89. this clearly shows the liver is infected with stage 4 fibrosis. This advanced liver fibrosis is due to cirrhosis. The radiology images shown in fig.3.1. (c) and fig.3.1. (d) are correlating with the clinical findings. And specifically, Prothrombin (PT) , international normalized ratio (INR) and Albumin are better readings to interpret the Stage of hepatitis. Serum albumin is significant characteristic parameter to interpret the progression of cirrhosis and fibrosis. Patients with significant decrease in serum albumin is specific indication of acute -on- chronic liver failure (ACLF). This chronic dis-function of the liver is highlighting chronic liver inflammations.

Serum Albumin normal range is 3.5 to 5.0. More specifically less than 4 g/dl is remarkably low reading indicative of ACLF. Normal values of INR are 0.9 to 1.2. This value is elevated for stage 2, stage3, and stage 4 are 1.0, 1.29 respectively. With increasing 0.1 value of INR, the blood is becoming thinner. Albumin and platelet count are more significant to determine the liver fibrosis. APP (Albumin platelet product) is much significant to diagnose liver fibrosis and cirrhosis. APP = Albumin ×Platelets / 1000. BUN (blood urea Nitrogen) is a waste material produced by the liver. This product is filtered using kidneys.[8][9] A high value of BUN is an indicative of liver damage or kidney damage. 6-18 mg/dL is the normal range. Elevation of this value is the indication of liver or kidney damage. The BUN values for patient 3 and patient 4 are 26.05 and 29.60. patient 3 BUN value is elevated 44.4

Percentage. this value is elevated 64.44 percentage for patient 4. These readings are validating the failure state of the liver function
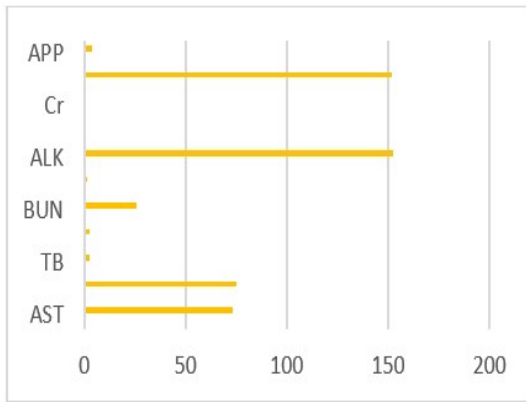

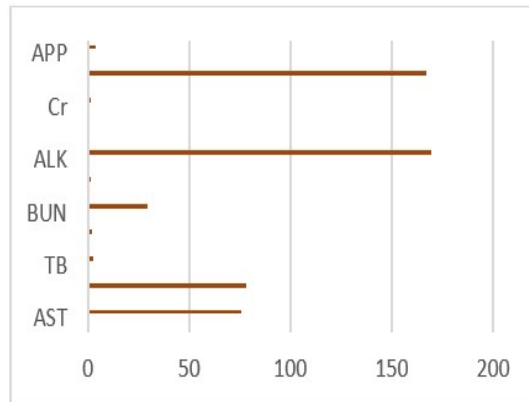Fig. 3.9 Liver Profile findings for Patient 3

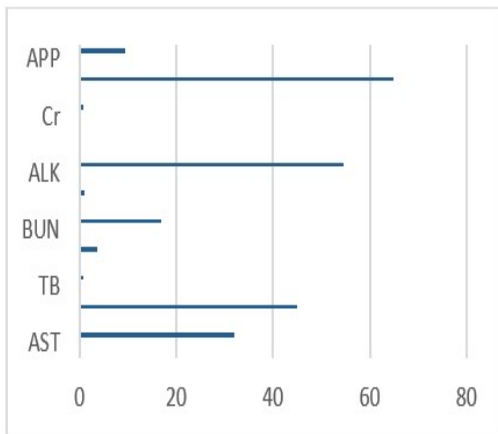Fig. 3.10 Liver Profile findings for Patient 4

Fig. 3.11 Liver Profile findings for Patient5

For liver cirrhosis APP value threshold value is 4.349. Low APP is due to chronic disorder of the liver. These clinical findings are correlated with the machine learning algorithms for meticulous detection of the liver fibrosis or cirrhosis with HEP-C. HCV (hepatitis -C Viral Load) test is another approach to detect the chronic infection of the liver. but the infection led to fibrosis or cirrhosis is difficult to interpret with HCV test. Machin learning algorithms augment with CT images improves the liver fibrosis stage and cirrhosis meticulously

CNN is applied to classify the normal or infected with hepatitis-C (Hep-C) or infected with fibrosis, or acute infection with cirrhosis. Fig.3.1(a) represents the non-infected liver images those images are normal images. The surface of the liver is smooth and regular contours of the liver is present. These clinical findings are correlated with the machine interpreted results.[10] The set of images shown in fig.3.1. (a) are normal images. Fig.3.1. (b) Represents the liver is infected with hepatitis-C viral infection. The machine interpreted images are correlated with clinical findings. The acute hepatitis represents liver hepatomegaly (enlarged liver) and periportal edema is traced. The heavily swollen liver is the serious inflammation led to adverse effects.[11]

Fig.3.1. (c) represents the liver is infected with Hepatitis-C and involved fibrosis. Right haptic lobe distance (RLD) value is 62mm and caudate lobe distance (CLD) value is 63mm. the caudate right lobe ratio is represented with the ratio of CLD/ RLD. [12]. The caudate right lobe ratio CRL-R is 1.016. The threshold value is 0.98. the caudate right lobe ratio is greater than 0.98. The clinical findings are correlated with machine findings. The fig.3.1. (c) shows fibrosis in liver.

Extensive scarring in the liver is pronounced significantly.

Fig.3.1. (d) represents the liver is infected with Hepatitis-C and traced cirrhosis. Clinical findings show significant change in density of spleen and homogeneous enhancement of liver. The right hepatic veins (RHV) found are 4.6mm right and the actual value is <5mm so is the evidence of cirrhosis. Clinical findings shows that main portal valve is 13mm size this is also significant evidence of cirrhosis. Irreversible scarring of large quantity of tissue traced, which is the indication of liver cirrhosis. This paper addressed the both clinical findings and machine learning algorithms to interpret the hep-C images. The research is focusing on deep machine learning algorithms for better interpretation of the hep-C images. Different classification algorithms were applied to interpret the images. This paper analysed clinical parameters to differentiate distinctive stages of the fibrosis meticulously. And these findings are corelated with machine learning algorithm.

4 CONCLUSIONS

The input CT image contrast is enhanced with contrast limited adaptive histogram equalization (CLAHE). The qualitative parameters like AMBE, PSNR, MSE, MD and MAE parameters are estimated. Low AMBE value is achieved while adjusting the contrast level of the input image. The PSNR value is $\geq$ 46 dB for all the processed images. The mean square error value (MSE) is less than '1' for all images. KNN methodology produced reliable results to distinguish the type of liver decease and its stage[14]. These findings are corelated with the clinical liver profile readings. The elevated levels of AST and ALT are significant representation of liver fibrosis and cirrhosis. 75.43 AST value and 78.34 ALT value show for patient '4 is the representation of stage 4 liver fibrosis and cirrhosis. Total bilirubin, albumin, alkaline phosphate, international normalised ratio is analysed to determine the level and type of disease. These clinical findings are coincident with the results extracted with proposed machine learning algorithm. Meticulous estimation is significant with CT images. Right haptic lobe

distance (RLD) value and caudate lobe distance (CLD) value are also investigated to strengthen the results with proposed algorithm. The caudate right lobe ratio is represented with the ratio of CLD/RLD. The caudate right lobe ratio CRL-R is 1.016. The threshold value is 0.98. The caudate right lobe ratio is greater than 0.98 indicate the liver is severely infected with bridging fibrosis with distortion and cirrhosis. The accuracy is superior while correlating the predicted results with the clinical findings [15][16]. This helps to give the result with better accuracy. which helps the medical field to identify the disease or the state of disease swiftly. The machine classification results are correlated with clinical findings. This paper focused to estimate the type of disease meticulously using KNN classification. The analysis is done both at more clinical findings and with machine learning algorithm which enhances the strength of the paper. The proposed methodology strengthens the image interpretation to detect the stage of liver fibrosis and hep -C infection.

REFERENCES

[1]. Akella, Aravind, and Sudheer Akella. "Applying machine learning to evaluate for fibrosis in chronic hepatitis c". medRxiv (2020).

[2]. Nandipati, Satish CR, Chew XinYing, and Khaw Khai Wah. "Hepatitis C virus (HCV) prediction by machine learning techniques" Applications of Modelling and Simulation 4 (2020): 89-100.

[3]. M. A. Khan et.al "Diagnosing the Stage of Hepatitis C Using Machine Learning" Journal of Healthcare Engineering, Volume 2021, Article ID 8062410, 10 December 2021.

[4]. Huber, Adrian, et al. "State-of-the-art imaging of liver fibrosis and cirrhosis: A comprehensive review of current applications and future perspectives". European journal of radiology open 2 (2015): 90-100.

[5]. Haga, Hiroaki, et al. "A machine learning-based treatment prediction model using whole genome variants of hepatitis C virus". Plos one 15.11 (2020).

[6]. Safdari, Reza, et al. "Applying data mining techniques to classify patients with suspected hepatitis C virus infection".Intelligent Medicine (2022).

[7]. Shergill, Annie, V. Lakshmi Narasimhan, and Emmanuella Budu. "S1112 Artificial Intelligence and Machine Learning-Based System for Diagnosis of Chronic Hepatitis C". Official journal of the American College of Gastroenterology| ACG 116 (2021).

[8]. Haga, Hiroaki, et al. "A machine learning-based treatment prediction model using whole genome variants of hepatitis C virus". Plos one 15.11 (2020).

[9]. Ioannou, George N., et al. "Assessment of a deep learning model to predict hepatocellular carcinoma in patients with hepatitis C cirrhosis.". JAMA network open 3.9 (2020).

[10]. Liu, Wei, et al. "Artificial intelligence for hepatitis evaluation" World Journal of Gastroenterology 27.34 (2021).

[11] Moradpour, Darius, François Penin, and Charles M. Rice. "Replication of hepatitis C virus". Nature reviews microbiology 5.6 (2007): 453-463.

[12] Azmi, Ahmad Najib, Soek - Siam Tan, and Rosmawati Mohamed. "Hepatitis C and kidney disease: an overview and approach to management". World journal of hepatology 7.1 (2015): 78.

[13] Surya Prasada Rao Borra et al.,"Pragmatic Investigations To Smart Dusts Location Appraisal Precisely Using Machine Learning", Journal of Theoretical and Applied Information Technology, 2023, 101(16), pp. 6301–6309.

[14] Surya Prasada Rao et. al.,"Unhealthy Plant Region Detection in Plant Leaves Using Adaptive Genetic Algorithms", Journal of Theoretical and Applied Information Technology, 2023, 101(8), pp. 3133–3137.

[15] K. V. Sagar et. Al.,"Saline fluid flow Supervision in Intensive Care Unit using Precision Algorithm", Journal of Theoretical and Applied Information Technology, 2023, 101(23), pp. 7769–7775.

[16] "An Improved CHI^2 Feature Selection based a Two-Stage Prediction of Comorbid Cancer Patient Survivability", Revue d'Intelligence Artificielle, 2023, 37(1), pp. 83–92.