Computer Methods and Programs in Biomedicine 155 (2018) 165–177

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Symtosis: A liver ultrasound tissue characterization and risk stratification in optimized deep learning paradigm

Mainak Biswas[a], Venkatanareshbabu Kuppili[a], Damodar Reddy Edla[a], Harman S. Suri[b,c], Luca Saba[d], Rui Tato Marinhoe[e], J. Miguel Sanches[f], Jasjit S. Suri[g,*]

[a] Department of Computer Science and Engineering, NIT, Goa, India
[b] Brown University, Providence, RI, USA
[c] Monitoring and Diagnostic Division, AtheroPoint™, Roseville, CA, USA
[d] Department of Radiology, A.O.U., Italy
[e] Liver Unit, Department of Gastroenterology and Hepatology, Hospital de Santa Maria, Medical School of Lisbon, Lisbon 1629-049, Portugal
[f] Bioengineering Department, IST, University of Lisbon, Portugal
[g] Advanced Knowledge Engineering Center, Global Biomedical Technologies, Inc., Roseville, CA, USA

ARTICLE INFO

Article history:
Received 1 July 2017
Revised 11 November 2017
Accepted 12 December 2017

1. Introduction

In the past two decades, liver diseases such as Fatty Liver Disease (FLD) are repeatedly listed as leading causes of liver related death in US among adults aged 45–54 years [1]. The deposition of fat in liver cells is called steatosis. Metabolic syndrome, consumption of alcohol, and obesity due to insulin resistance are some of the causes for FLD [1,2]. NAFLD is a frequent cause of chronic liver diseases, making up 19–46% of all liver diseases in the western world [2].

Several studies have been conducted for FLD tissue characterization. Lamb et al., [3] used dual energy CT (DECT) based technique for classification Computed Tomography (CT) liver images. Guo et al. [4] used neural networks for classification of Magnetic Resonance Imaging (MRI) liver images. Suri et al. [5] designed a computer aided liver data classifier for Ultrasound in 2012. They used Decision Tree [6], a machine learning (ML) method for the classification leading to an accuracy of 93.3%. In 2016, under the same class of Symtosis™ (Global Biomedical Technologies, Inc., Roseville, CA, USA), Suri et al. [7] proposed Levenberg-Marquardt back propagation neural network (BPNN) [8] classifier, that gave an accuracy of 97.6%. Often, the FLD US images are hypoechoic preventing statistical classifiers to reach 100% accuracy. One of the attributing causes is the lack of appropriate tissue characterization features for the varying disease characteristics. Further, the cross-validation protocols require several trials per combination during training, proving to be time consuming. Finally, the tissue-based feature maps are applied to entire region of interest picking up erroneous maps which lowers the efficiency of the ML systems.

Convolution neural networks recently became active in classification framework [9]. An example of CNN model can be seen in Fig. 1. In CNN, a series of convolutions (top purple triangles pointing south), rectilinear unit (ReLu) and pooling operations (bottom purple triangles pointing north) are applied for feature detection. A convolution operation is a weighted average operation of a convolution filter with an input image. A ReLu, which is often performed with convolution, performs the task of rectification of outputs of convolution to reduce the likelihood of the gradient to vanish. Pooling is done for down sampling of input data. We have started to see the role of CNN in several worldly applications buffered with optimization techniques [10].

In this paper, we applied a Deep Learning (DL) technique for detection of hypoechoic FLD and stratification of normal and abnormal US liver images under the class of Symtosis™. In here, our focus is on characterization of US liver images as normal or abnormal (FLD afflicted). This is the first study of its kind that presents a foremost comprehensive comparison of three ML-based classification methodologies: namely, support vector machines (SVM) [11], extreme learning machines (ELM) [12] and DL CNN model. Further, in this study, we additionally investigate a specialized operation called inception. The inception operation integrates different types of convolutions using dimensionality reduction principle followed

* Corresponding Author.
E-mail addresses: jsuri@comcast.net, jasjit.suri@atheropoint.com (J.S. Suri).

https://doi.org/10.1016/j.cmpb.2017.12.016

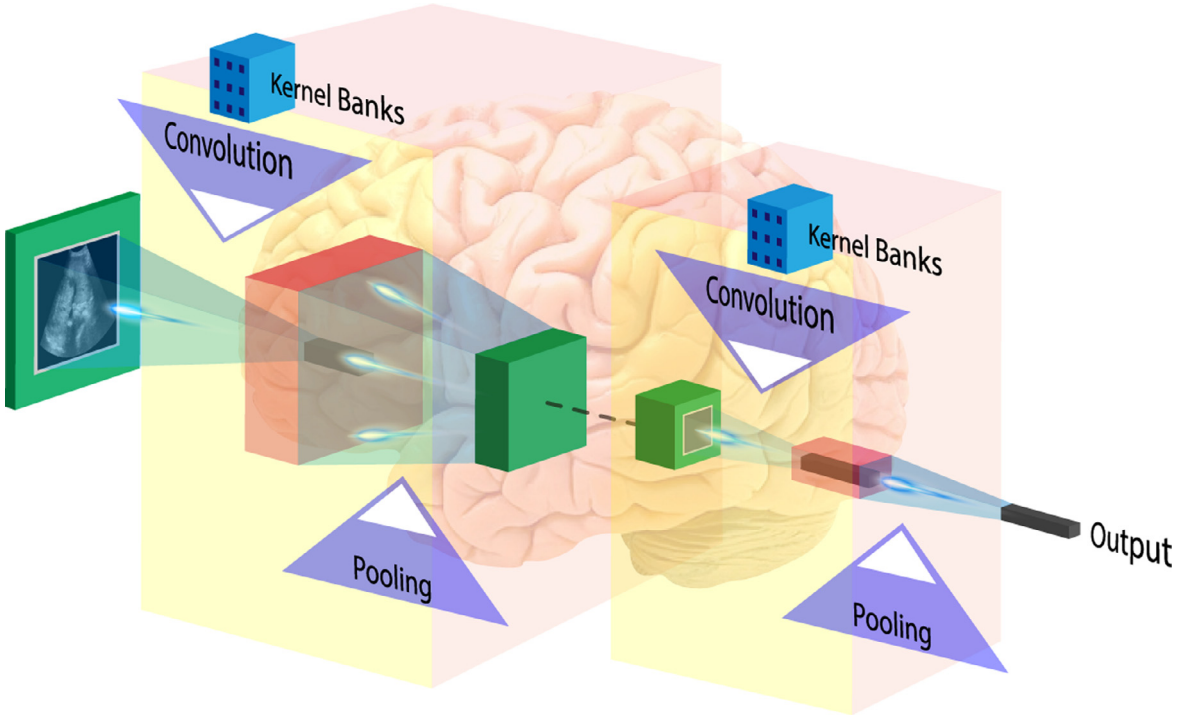

Fig 1. The basic Deep Learning CNN model.

by concatenation of these different convolution pipelines thereby reducing the computational cost of the DL liver Symtosis^{TM} system. Inception leads to an increased depth of DL. The increased depth allows the system to extract higher level features resulting in greater accuracy of the system. Lastly, we explore GPU-based framework in our DL-based application for tissue characterization and risk assessment. The AUC obtained from ROC analysis for SVM, ELM and DL are: 0.79, 0.92 and 1.0 leading to perfect accuracy of 100% using the *K10* cross-validation protocol for optimized US liver images. The global architecture of the DL model is given in Fig. 2. The US liver images are border stripped optimized for best DL stratification accuracy which becomes the basis for all data analysis. Further, our system has been validated using biometric facial data.

The layout of the paper is as follows. Section 2 describes patient demographics and acquisition while Section 3 discusses methodology for DL-based Symtosis^{TM} system and two other conventional classification algorithms. The results, performance evaluation and discussion are presented in Sections 4 –6, respectively. Finally, our proposed study concludes in section 7.

2. Patient demographics and acquisition

A total of 63 patients were selected from the Gastroenterology Department of the Santa Maria Hospital (ethics approval granted), in Lisbon, Portugal [13], also used in our previous study [5,7]. The US images were acquired using CX_c 50 (Philips Medical Systems) in DICOM (Digital Imaging and Communications in Medicine) format. The scanner images were 8 bits a pixel resolution consisting of 1024×1024 sizes. The US scanner was equipped with curved array transducer C5-1. This consisted of 160 piezoelectric elements with a resonant frequency of 1–5 MHz. Out of the 63 patients, 36 patients were diagnosed with FLD and rest 27 patients were diagnosed as normal. The US images of normal and fatty livers were acquired by expert operators with US equipment in the hospital facility. The resulting images were annotated as normal and abnormal depending upon the indicators obtained from the biopsy reports from laboratory. These annotated images formed the Ground Truth (GT) or gold standard for our experiment.

The liver has a small left lobe (in the epigastric area) and a large right lobe (in the right hypochondrium. The right liver is the major liver part, we use scanned image of the right lobe liver. In the US image of a normal liver, the echogenic intensity of liver parenchyma and kidney cortex is similar whereas in the US image of a steatotic liver, there is an increased echogenic intensity of the liver parenchyma which is clearly brighter than the kidney cortex [14,15]. A region of interest (ROI) of 128×128 pixels along the medial axis was extracted from each image. The original liver images are shown in Fig. 3. The top row shows 4 normal and bottom row shows 4 abnormal images consisting of hypoechoic and hyperechoic tissue regions.

3. Methodology

The basic architecture of DL using CNN was earlier shown in Fig. 1, where low level features (deck of cards) were converted into high level features (single layered, interpretable by humans) during the neural network (NN)-based learning process [16]. CNNs apply multiple layers of convolution and pooling to learn this deck of features derived from original image. The first challenge in a basic CNN model is the memorization of the NN leading to over fitting the input data and lowering the classification accuracy. This is due to a combination of non-optimized number of hidden layers, type of cross-validation protocol and type of convolution deck created. The second challenge arises due to the requirement of more number of layers, since these layers provide the power to the design of the feature maps which in turn leads to separability of the classes. However, this increases the computational complexity of the DL system. The solution to the first challenge is the introduction of a "dropout" strategy, thereby using a limited number of hidden layers (weights) and this can be adapted randomly. The solution to the second problem is introduction of an "inception model" whose ultimate role is to reduce the computational cost. These two architectural inclusions to the basic CNN form our model for liver dis-

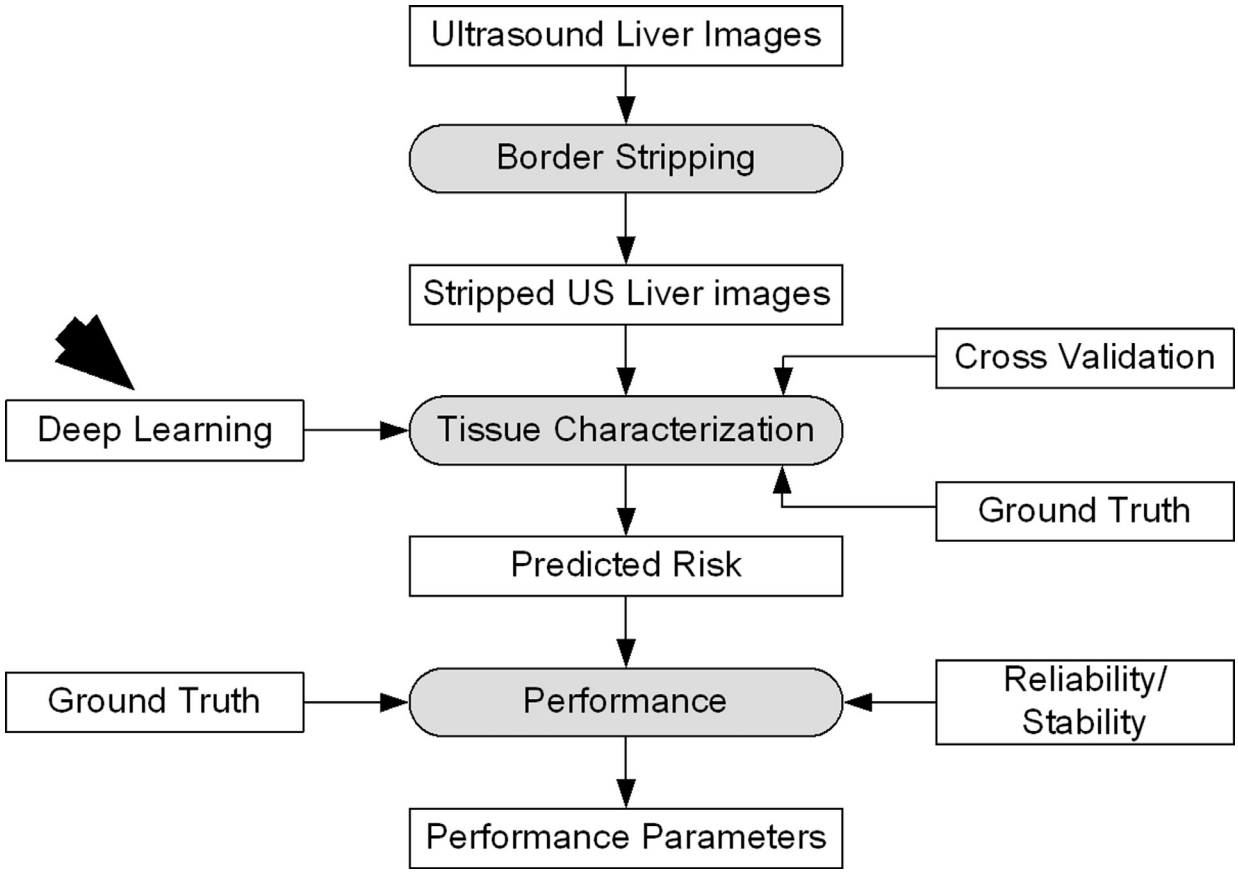

Fig 2. DL-based Symtosis™ system.

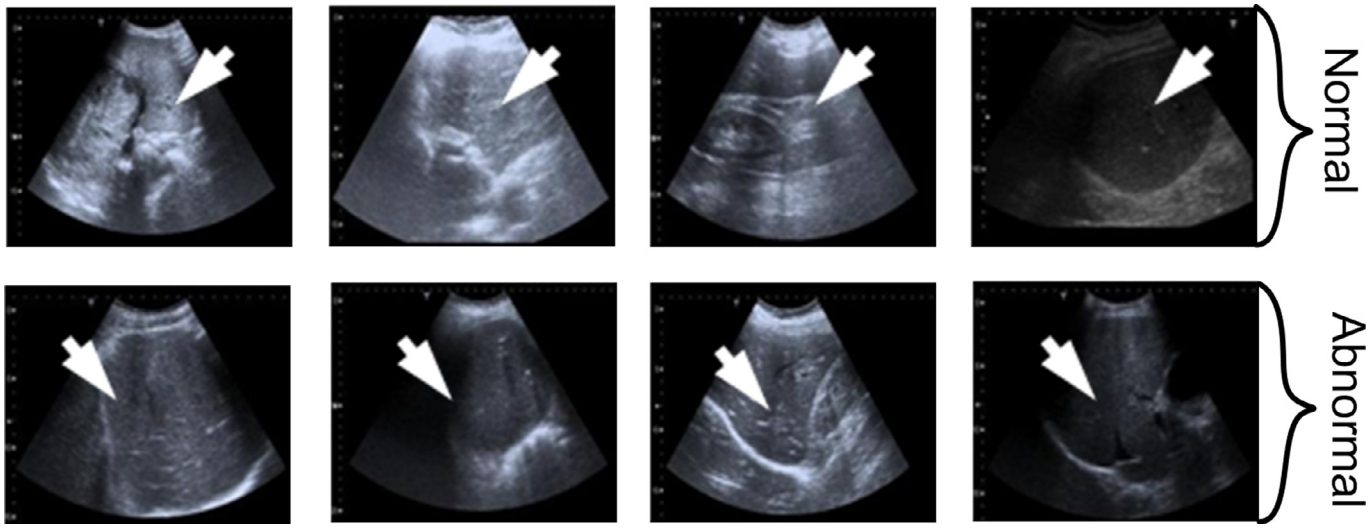

Fig 3. Normal liver images (top row) and abnormal liver images (bottom row).

ease stratification provides the fundamental novelty of our system. This architecture is shown in Fig. 4. This DL architecture shows two sets of convolution, nine inceptions, five pooling, one dropout, one linear and one softmax (binary classifier) cascaded layer.

3.1. Risk stratification model using tissue characterization

The basic model for risk stratification can be characterized using the conventional approaches of the literature as developed by Suri and his team and has been applied to lung [17], liver [5], prostrate [18] or ovarian [19] diseases.

It is observed that FLD images show both kinds of echogenic intensities: hypo and hyper, thus making them harder to detect and stratify. The most fundamental difference between our current approach and the previous machine learning paradigms is the adaptation of DL-based feature extraction, feature reduction and training of large number of NN-weights (over the order of one million), unlike the conventional approach, that uses support vector machine (SVM) for training the derived greyscale features. This can

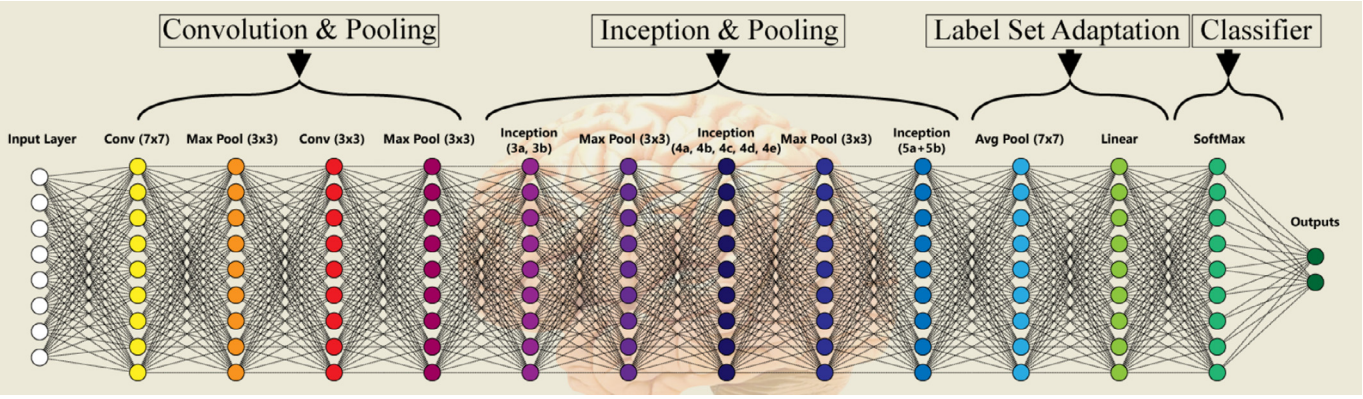

Fig 4. The CNN model showing convolution, pooling, inception, linear and softmax layers.

be seen in Fig. 5. To be more explicit, since DL requires a large number of features due to convolutions, one requires operating the process of embedded feature extraction (due to convolution, pooling and ReLu) in the region of interest that purely contains the tissue information and no redundant background non-tissue region or noise.

We thus use a stripping protocol that removes the background information. Followed by this cropping procedure, the current paradigm uses the same cross-validation (CV) protocol (so-called *K*-fold cross validations: $K = 2$, 3, 5 and 10) as conventional machine learning system does. *K2* cross-validation protocol divides the data into two parts. There are two rounds for training/testing for each combination that is, one part for training and other part for testing. Similarly, *K3* divides the data into three parts say A, B and C. One part among them is selected for testing and other two parts are combined for training. Likewise in *K2*, three rounds of training/testing takes place for each combination. This same concept is applied for *K5* and *K10* cross validation. . Suri's group has spanned application of Machine Learning in different organs such as Carotid [20], Coronary [21], Liver [7,22], Breast [23,24], and Prostate [25]. No matter how many types of classes (or stages or grades) of liver disease exists, it is definitely possible to classify such disease types as long as the corresponding ground truth information for these class types (stages of liver diseases) is available. This was recently demonstrated by Suri's group using Bayesian Classifier for skin disease classification for four different types of classes adapted [26].

3.2. Building blocks of CNN

The fundamental building blocks of CNN architecture consists of: convolution, pooling, rectified linear unit (ReLu), dropout, fully connected layer and softmax (already presented in Fig. 4). In convolution, different convolution filters (kernels) are applied to original images to obtain features, also known as feature maps. Further discussion on CNN can found in Appendix A.1.

The DL architecture table is described in next subsection. The data flow within the system is shown in Fig. 6. It is a 22-layered (see Table 1, column labelled as Depth) DL architecture for classification of US liver images and this consist of: convolution, pooling, inception, dropout, linear and softmax. The US liver images are fed into the DL architecture for computing the training weights which are then fed to test data to obtain the final accuracy values using softmax layer. The description and of each layer of DL's data flow is given in Table 1. In the table, #1×1 refers to the number of 1×1 convolutions applied to the input for dimensionality reduction. Reduce # 3×3 by 1×1 convolution filter means application of 1×1 convolution filters prior to the application of number of 3×3 convolution filter. Similarly, it implies for "Reduce # 5×5". The reduce pool number refers to the number of 1×1 convolutions applied after pooling operations for dimensionality reduction. The last column shows the number of parameters or weights that are to be trained. The application of 1×1 convolutions for dimensionality reduction in a layer increases its depth by one. The application of inception within the DL architecture allows integration of convolution filters of different sizes into one single inception layer. A brief description of DL architecture can be found in Appendix A.2.

4. Results

The results of the risk stratification can be characterized by evaluating the final accuracy of the DL system and comparing it to other conventional machine learning systems (SVM and ELM) using the class of Symtosis™ system discussed in Fig. 2. There are two components to be demonstrated in the results: (i) optimization of the image size for background removal and (ii) understanding the effect of the data size on the three learning methods.

4.1. Image optimization for three learning methods

It is critical to ensure that we remove the background information prior to estimating the accuracy on the test data using the cross-validation approach as presented in Fig. 5. Border stripping was therefore applied to the US liver images to remove the background ensuring that the region of interest only has the tissue region. Feeding these cropped images to the Symtosis™ system and running the SVM/ELM/DL architectures using *K10* CV protocol, we observe the stratification accuracies as: SVM: 82%, ELM: 92% and DL: 100%. The optimization curve between percentages of border stripping vs. accuracy for 63 patients is shown in Fig. 7. It shows highest accuracy for 15% stripping of the background region corresponding to *K10* cross-validation protocol. The analogous accuracy bar chart is given in Fig. 8. Our observation showed that SVM and ELM is less sensitive to background information compared to DL. One possible reason is over-fitting in the DL scenario unlike SVM and ELM that reaches its optimal values in the first run and no iterations are involved. Second possible justification is quick generalization and retaining that generalization, unlike DL that first memorizes and then falls to generalization. The corresponding bar chart for 15% stripped images for all three strategies can be seen in Fig. 8, marked as (a), (b) and (c), correspondingly for SVM, ELM and DL learning architecture. The average accuracy, AUC, sensitivity, specificity and PPV values for *K10* cross validation for 15% stripped US liver images are shown in Table 2. The results show 100% accuracy, sensitivity, specificity and PPV values for DL, which is better than SVM and ELM. The AUC value is perfect 1.0 which is

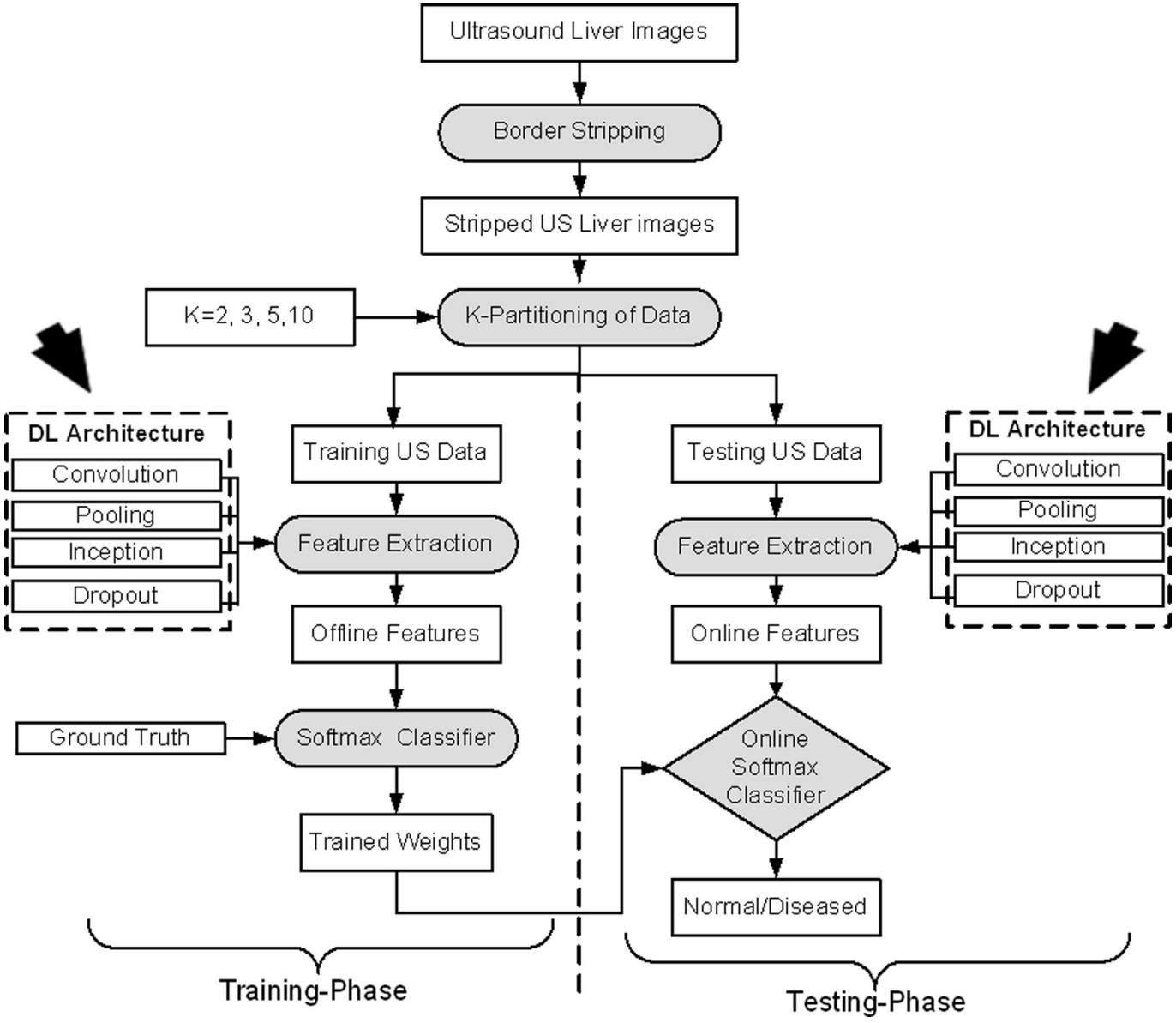

Fig 5. DL tissue characterization system using Symtosis™.

Table 1
DL architecture table showing different layers.

SN*	Type	Patch size/ stride	Output Size	Depth	#** 1 × 1	Reduce #3 × 3 CF* by 1 × 1CF	# 3 × 3 Conv.	Reduce #5 × 5 CF by 1 × 1CF	#** 5 × 5 Conv.	Reduce Pool	Number of Weights
1	Convolution	7 × 7/2	112 × 112 × 64	1	–	–	–	–	–	–	2.7 × 10^3
2	Max Pool	3 × 3/2	56 × 56 × 64	0	–	–	–	–	–	–	
3	Convolution	3 × 3/1	56 × 56 × 192	2	–	64	192	–	–	–	112 × 10^3
4	Max Pool	3 × 3/2	28 × 28 × 192	0	–	–	–	–	–	–	–
5	Inception (3a)	–	28 × 28 × 256	2	64	96	128	16	32	32	159 × 10^3
6	Inception (3b)	–	28 × 28 × 480	2	128	128	192	32	96	64	380 × 10^3
7	Max Pool	3 × 3/2	14 × 14 × 480	0	–	–	–	–	–	–	–
8	Inception (4a)	–	14 × 14 × 512	2	192	96	208	16	48	64	364 × 10^3
9	Inception (4b)	–	14 × 14 × 512	2	160	112	224	24	64	64	437 × 10^3
10	Inception (4c)	–	14 × 14 × 512	2	128	128	256	24	64	64	463 × 10^3
11	Inception (4d)	–	14 × 14 × 528	2	112	144	288	32	64	64	580 × 10^3
12	Inception (4e)	–	14 × 14 × 832	2	256	160	320	32	128	128	840 × 10^3
13	Max Pool	3 × 3/2	7 × 7 × 832	0	–	–	–	–	–	–	–
14	Inception (5a)	–	7 × 7 × 832	2	256	160	320	32	128	128	1072 × 10^3
15	Inception (5b)	–	7 × 7 × 1024	2	384	192	384	48	128	128	1388 × 10^3
16	Avg. Pool	7 × 7/1	1 × 1 × 1024	0	–	–	–	–	–	–	–
17	Dropout	−40%	1 × 1 × 1024	0	–	–	–	–	–	–	–
18	Linear	–	1 × 1 × 1000	1	–	–	–	–	–	–	1000 × 10^3
19	Softmax	–	1 × 1 × 2	0	–	–	–	–	–	–	–

*SN: Serial Number; **#: Number of; CF: convolution filter

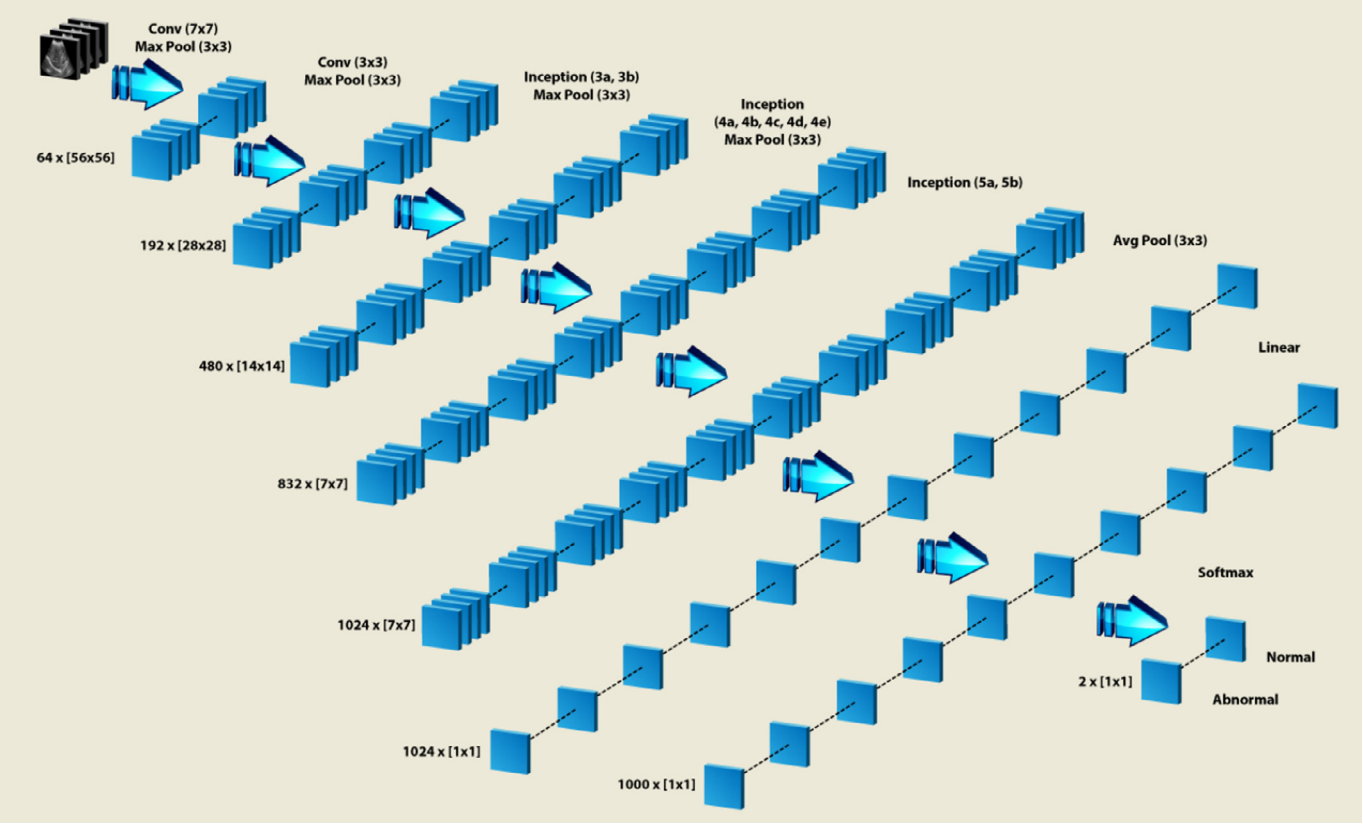

Fig 6. Data flow in DL architecture.

Table 2
Comparative performance of three learning architectures using *K10* cross-validation protocol: SVM vs. ELM vs. DL.

Classifier Type	Average ACC* (%)	Average Sensitivity (%)	Average Specificity (%)	Average PPV (%)	Average AUC
SVM	82.08	64.21	93.56	86.31	0.79
ELM	92.22	93.33	90.83	84.58	0.92
DL	100.00	100.00	100.00	100.00	1.00

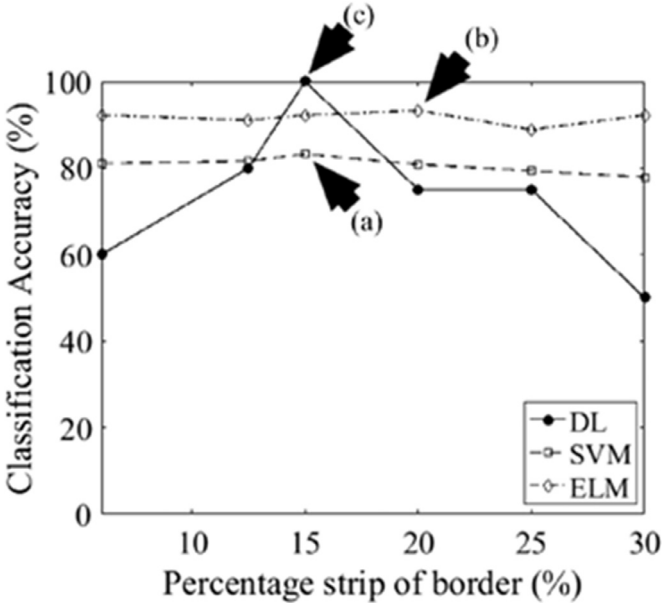

Fig 7. Accuracy vs. percentage strip of background borders in liver scans. The accuracy was computed using *K10* protocol in SVM (a), ELM (b) and DL (c) frameworks.

greater than SVM and ELM values. It is observed from our experiments that Deep Learning shows robust performance to noise during the cropping protocol. The cropping function is applied to remove background noise information. It is seen that Deep Learning shows 100% accuracy at 15% cropping of border zones. However, a further increase in cropping removes FLD information from liver which results in drop of accuracy leading to optimized cropping value. These further shows that tissue information of FLD is not always centralized in images. Further, we think the inception model really yields effective results when the border cropping region is 15%. This is due to the same rationale that tissue is best represented here and noise characteristics are least. Note that these advanced features do not exist in the conventional machine learning system like SVM or ELM, however the optimized values are depicted with less prominence.

4.2. Effect of data size on stratification accuracy using K*-fold cross-validations*

The effect of data size during the training and testing is important to understand in the DL framework and to compare against the SVM and ELM strategies. Using the four set of cross-validation sets (*K2, K3, K5*, and *K10*), the Symtosis^{TM} system was executed by following the strategy of Fig. 5. The results are shown in Fig. 9. As can be seen in the figure, the DL out performs SVM and ELM architectures in terms of accuracy.

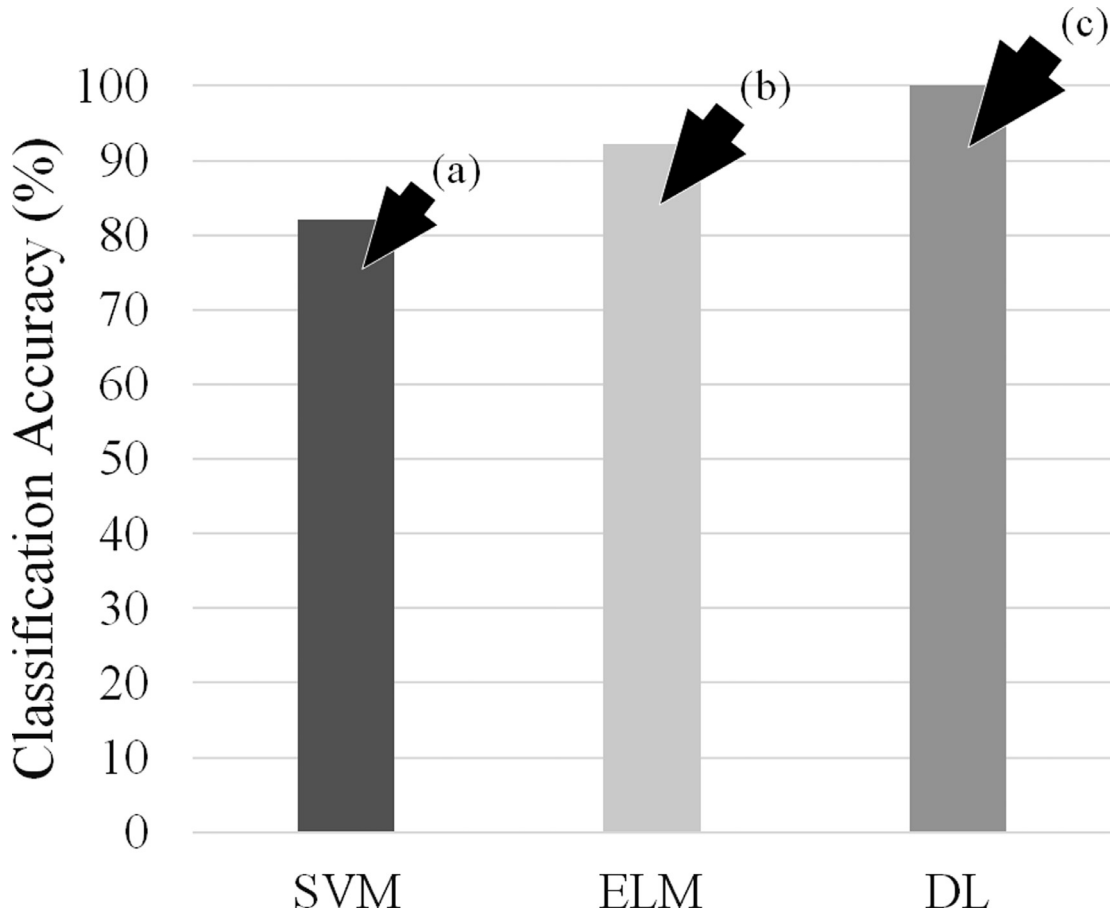

Fig 8. Bar chart of optimized accuracies for SVM, ELM and DL systems using *K10* protocol.

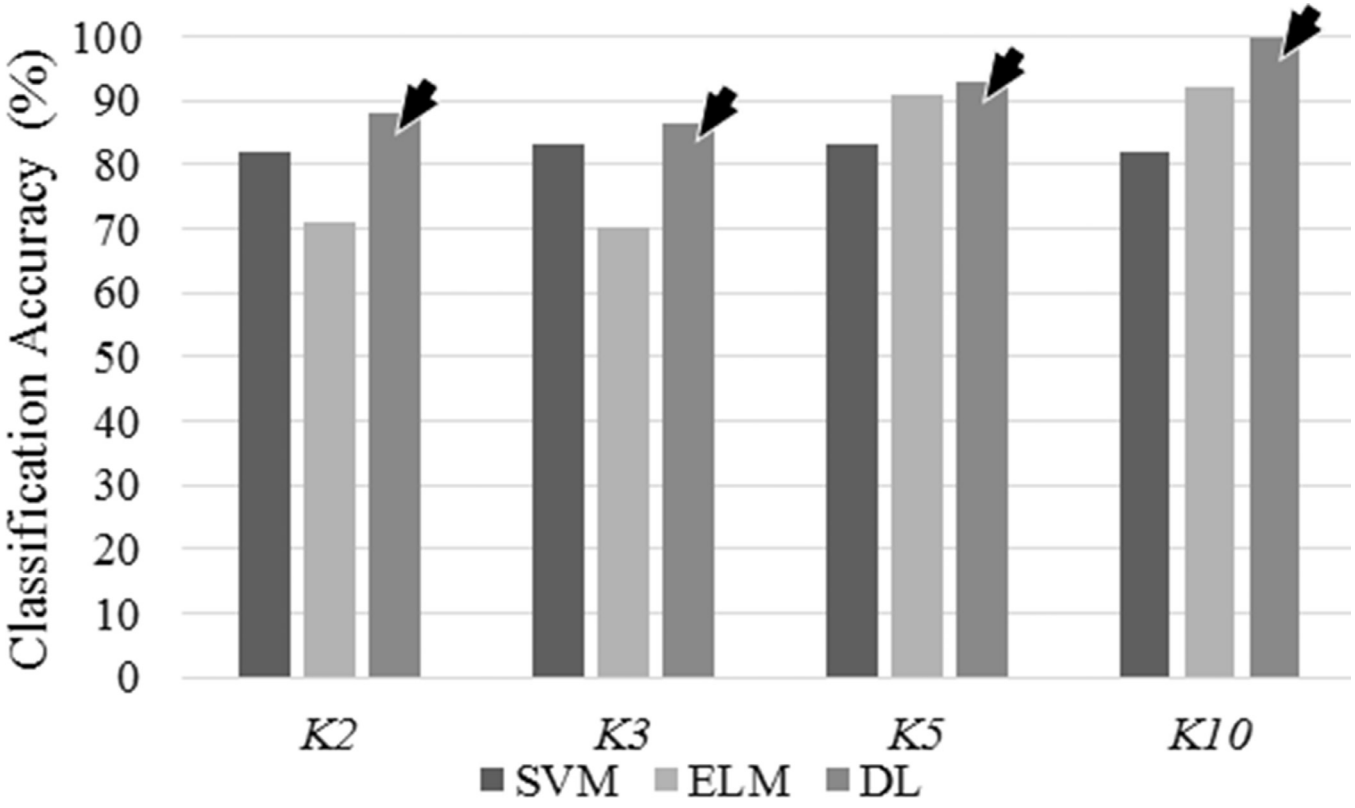

Fig 9. Bar chart for accuracy vs. type of CV protocol (*K2*, *K3*, *K5* and *K10*) for three sets of risk stratification strategies (SVM, ELM and DL).

Table 3
Mean feature values for normal and abnormal liver patients.

ELM/SVM			DL		
Abnormal	Mean Abnormal Liver Values	SD* Abnormal	Abnormal	Mean Abnormal Liver Values	SD Abnormal
Gabor	0.5000	0.5063	Deep Features	0.2593	0.2469
GLRM	0.2954	0.2584			
GLCM	0.4860	0.2865			
Normal	Mean Normal Liver Values	SD Normal	Normal	Mean Normal Liver Values	SD Normal
Gabor	0.4617	0.3799	Deep Features	0.2849	0.2462
GLRM	0.3214	0.3962			
GLCM	0.5037	0.2839			

*SD: Standard Deviation.

Table 4
Liver segregation index.

Classifier	Liver Segregation Index (LSI) (%)		
	Gabor	GLRM	GLCM
ELM SVM	7.65	8.81	3.66
DL	Deep Learning Features		
	9.86		

4.3. Stratification analysis using "Liver Segregation Index"

It is important to understand how Symtosis™ class of features behave in the three sets of paradigms, i.e., SVM, ELM and DL. The basic idea here to compute the features, normalize them and compute the index which can appreciate the difference between the normal and abnormal liver patients. We defined this index as: Liver Segregation Index (LSI) and is mathematically given as: $\mathrm{LSI} = \frac{|\mu_{Nor} - \mu_{Abnor}|}{\mu_{Nor}} \times 100$, where μ_{Nor} and μ_{Abnor} are the mean of normal and abnormal grayscale tissue features and is given as:

$$\mu_{Nor} = \frac{\sum_{j=1}^{P_{Nor}} \sum_{i=1}^{N} x_i^j}{P_{Nor} \times N} \; ; \; \mu_{Abnor} = \frac{\sum_{j=1}^{P_{Abnor}} \sum_{i=1}^{N} x_i^j}{P_{Abnor} \times N} \tag{1}$$

where, P_{Nor} and P_{Abnor} indicates total number of normal and abnormal patients, where N represents the total number of features. The results of the mean value of the features using three strategies for the normal and abnormal classes are given in Table 3. The corresponding LSI values are shown in Table 4. Note that there is there is a significance difference between DL vs. ELM/SVM architectures. Note that SVM/ELM uses Gabor, GLRM and GLCM feature maps [5]. The details of Gabor, GLRM and GLCM is given in Appendix C. This separation in the feature map is another justification of higher accuracy in DL compared to ELM and SVM.

It is clearly seen from Table 4 that deep feature LSI values are higher than ML feature LSI values. The plot for LSI (%) vs. deep features is given in Fig. 10 (i). It is seen that LSI values increases gradually (arrow (a)) till the total number features reaches 800 (arrow (b)) and thereafter increases sharply (arrow (c)).

5. Performance evaluation: ROC, reliability and timing analysis

The DL architecture is evaluated using K2, K3, K5 and K10 cross-validations protocols against SVM and ELM. In Section 5.1 we discuss ROC analysis. Further, the Section 5.2 shows reliability analysis for different K-fold cross-validations and Section 5.3 presents timing analysis.

5.1. ROC analysis

The ROC curve for SVM, ELM and DL is shown in Fig. 10(ii). The ROC curve for SVM, ELM and DL are shown as (a), (b) and (c). Our observations show that area under the curve (AUC) for DL relatively improved compared to SVM and ELM.

5.2. Reliability analysis

Reliability Analysis is performed for SVM, ELM and DL for all K-fold cross validations. The reliability index $\zeta_{N_L}(\%)$ is formulated as:

$$\zeta_{N_L}(\%) = \left(1 - \frac{\mu_{N_L}}{\sigma_{N_L}}\right) \times \; 100 \tag{2}$$

where, μ_{N_L} is the mean accuracy and σ_{N_L} represents the standard deviation of all accuracies for N_L liver images. The bar chart of reliability analysis is shown in Fig. 11. The chart clearly shows that the DL architecture is more reliable than SVM and ELM.

5.3. Timing analysis

The timing analysis of DL architecture with SVM and ELM is given in Table 5. It is seen that average training time and testing time of DL architecture is slower than SVM and ELM. Further, we compute the product of classification accuracy and testing time as PE-index for evaluating the composite performance (5th column in Table 5).

6. Discussion

The focus of this study was FLD or hepatic steatosis classification using a class of Symtosis™, a DL-based strategy along with benchmarking against SVM and ELM paradigms. The characterization of FLD has been attempted using conventional ML-based techniques such as SVM [11], BPNN [8], PNN [27], FC [28] and DT [6]. The hypoechoic nature of the liver FLD US images has always posed a challenge in detection and stratification by conventional ML-based methodologies. The ML techniques applied for FLD detection and risk stratification using US have limitations in computing feature maps thereby limiting the accuracy. Further, ML algorithms are suited for smaller data sets where dynamic fitting is not required [5,29–33]. In here DL was applied for FLD detection and stratification. The Symtosis™ system is validated using biometric facial dataset that gave an accuracy of [34,35] (see Appendix B).

6.1. Benchmarking

There are hardly any DL-based applications for liver detection and risk stratification. We therefore have decided to benchmark our DL-strategy with latest ML paradigms applied to US imaging. The results of our benchmarking are presented in Table 6. It is important to see that the ML strategies applied to US such as liver, thyroid, ovarian primarily came from couple of major groups. The crux in this table is the column 5 which shows the feature maps adapted in their ML-based techniques. These features were mainly

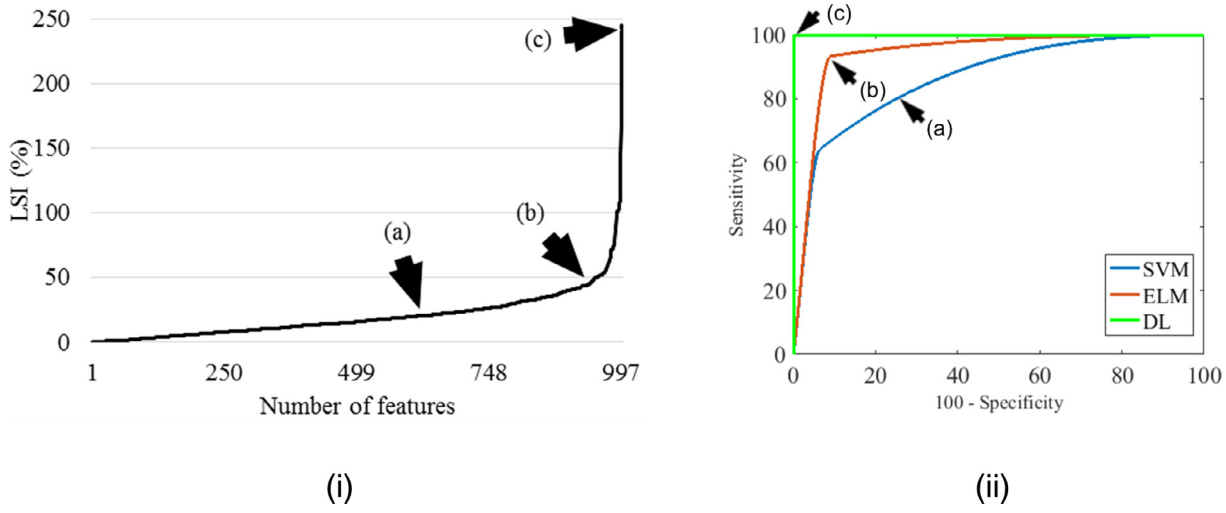

Fig 10. (i) LSI (%) vs. number of features using DL framework. (ii) ROC curve for SVM (a), ELM (b) and DL (c) ($K = 10$).

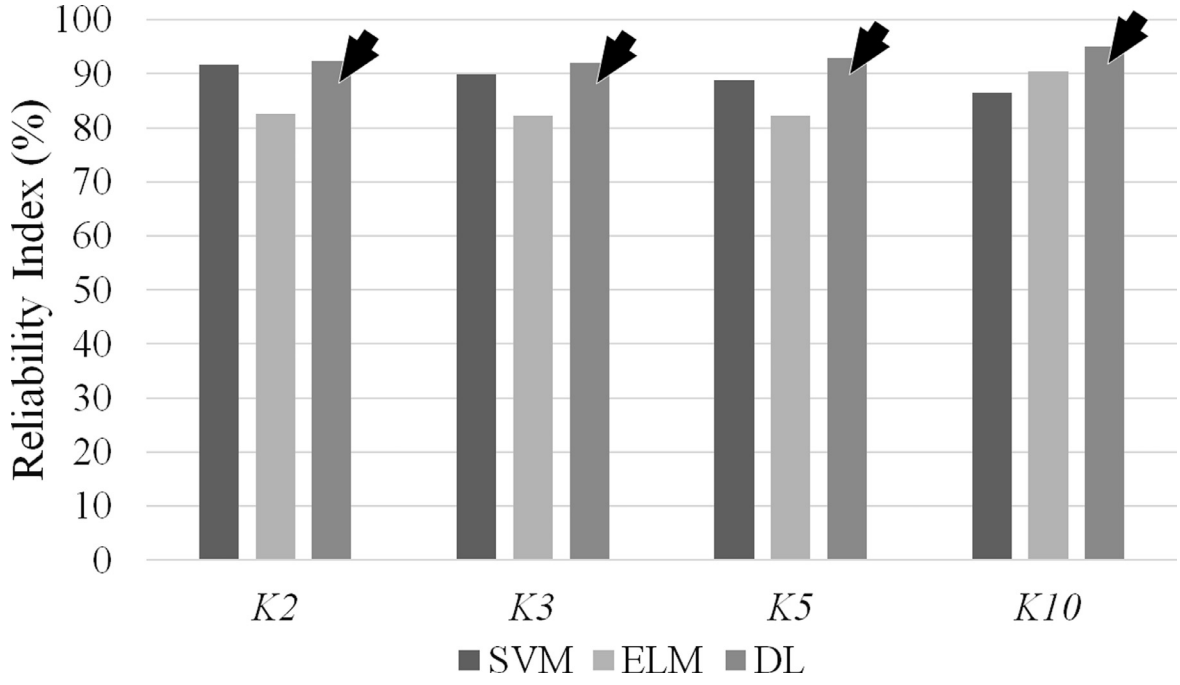

Fig 11. Bar chart for reliability index for different *K*-fold (*K2, K3, K5* and *K10*) cross validation schemes for three set (SVM, ELM and DL) of learning methods.

Table 5
Timing analysis.

SN	Type of Classifier	Avg. Train Time (ms*)	Avg. Test Time (ms)	PE-Index Test
1	SVM	0.029016	0.002262	0.188269
2	ELM	0.007020	0.001560	0.143868
3	DL	1.300937	0.002476	0.247600

*ms: milliseconds.

Table 6
Benchmarking table (in chronological order).

SN	Reference	Disease	Data set	No. of Samples	Feature Extraction	Type of Classifier	Accuracy (%)
1	Acharya et al. [5] (2012)	FLD*	liver US images	100	HOS*, Texture & DWT*	DT*	93.3
2	Acharya et al. [36] (2014)	HT*	thyroid US images	526	Wavelet Transform	FC*	84.6
3	Acharya et al. [19] (2014)	OC*	transvaginal US images	2600	HIM*, Gabor & Entropies	PNN*	99.8
4	Saba et al. [7] (2016)	FLD	liver US images	62	Harlick, Fourier & Gabor	BPNN*	97.6
5	Liu et al. [37] (2017)	LC*	liver US images	91	DET* + T-CNN*	SVM*	89.2
6	Proposed method	FLD	liver US images	63	CNN*	Softmax	100.0

*FLD: Fatty Liver Disease; LC: Liver Cirrhosis; OC: Ovarian Cancer; HT: Hashimoto Thyroiditis; CNN: Convolution Neural Networks; T-CNN: Trained CNN; DET: Detection of Liver Capsule; HIM: Hu's invariant moments; HOS: High Order Spectra; DWT: Wavelet Packet Decomposition; SVM: Support Vector Machine; BPNN: Back Propagation Neural Network; PNN: Probabilistic Neural Network; FC: Fuzzy Classifier; DT: Decision Tree

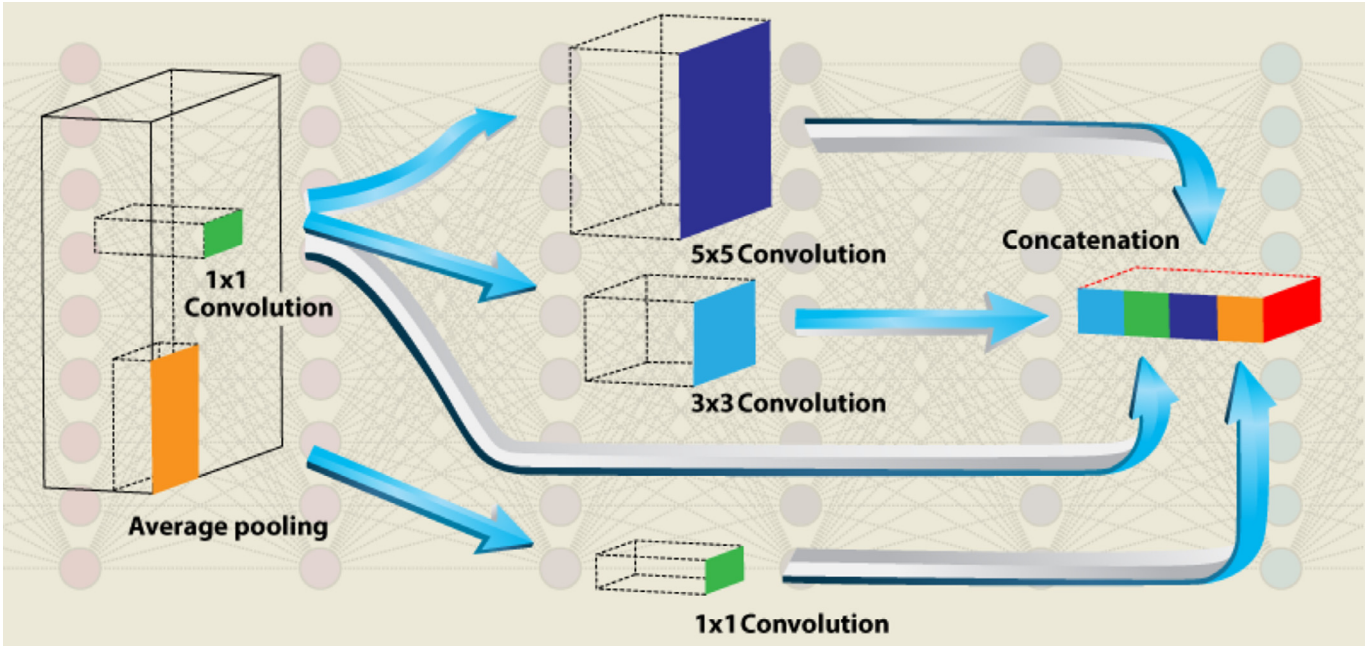

Fig 12. Inception layer showing dimensionality reduction and concatenation process utilized in the DL architecture.

texture, wavelet transform, Gabor, Fourier types and higher order spectra (HOS).

Suri and his team (in 2012) [5] achieved an accuracy of liver stratification of 93.3% using a combination of features such as: high order spectra, texture and wavelet packet decomposition feature while using decision tree (DT) classifier. A similar approach using fuzzy classifier and also using wavelet transform based feature was again attempted in 2014 by Suri and his team [36], and yielded an accuracy of 84.6% on thyroid US images. A closer application of Suri and his team's work [19] was using a neural network approach on ovarian US images using a combination of Hu's invariant moments, Gabor and entropy based features using probabilistic neural network classifier leading an accuracy of 99.8%. More recently, Suri and his team (2016) [7] presented a back propagation neural network classifier for liver US that used features such as Harlick, Fourier and Gabor leading to an accuracy of 97.6%. More recently, Liu et al. [37] used a combination of liver capsule detection technique and trained CNN model for feature extraction, and used SVM as classifier to achieve accuracy of 89.2%. We are in the same space as CNN with stronger and encouraging results nearly reaching 100%, even with a smaller data size of 63 subjects. The CNN showed high AUC value of 1.0 for *K10* cross validation as classification accuracy reached 100%.

6.2. Strengths and weakness

The power of decomposition of images into small patches and then extracting larger deck of features maps allows neural networks to learn, generalize and accurately stratify the FLD disease. This DL-based strategy can span the network from single layer to as large as 22 ensuring selection of refined features by computing millions of weights in dynamic our framework. Such a model provides inherent strength in classification paradigm. Such a model provides inherent strength in a classification paradigm. The application of inception layer as shown in Fig. 12 further strengthens the CNN model by concatenating several convolution, Relu and pooling operations in a single layer using the concept of dimensionality reduction (explained in Appendix A.3). This leads to increase in depth without increase in computational cost in the proposed DL model. Further due to dimensionality reduction combined with concatenation power of the inception model, when implemented in GPU framework, the DL-based strategy offers diverse applications and extendibility.

In spite of the above distinguishing and unique benefits, there are some challenges that needs to be addressed over time. There is always a price to pay for higher accuracy and large number of neural network layers. Since it is handling millions of shared weights, the systems takes longer convergence time unlike ELM, which only has single layer. The DL system is needed to be validated using transfer learning [38] i.e., training DL in one dataset and testing in another to reach generalization. Even though, our solution provides nearly 100% accuracy and results are encouraging, it requires more work for ensuring that these evolving architectures are stabilized, robust, fast and generalized.

7. Conclusion

We showed a DL-based Symtosis^{TM} system for liver ultrasound tissue characterization and risk stratification of normal and abnormal liver images which possesses both hyper- and hypoechoic intensities. Our DL-based Symtosis^{TM} system uses inception model for dimensionality reduction and speed of our DL network without increasing computational costs. The images have been optimized by removing background from the original liver images by stripping the border. Our results demonstrate a flawless accuracy by using 15% background removal. The liver data set has been cross-validated (*K2, K3, K5* and *K10* protocols) and compared with conventional ML techniques such as: SVM and ELM. Further, the results from our experiment shows that our DL-based Symtosis^{TM} system is reliable and stable compared to SVM and ELM across all cross validations. The features extracted from ML techniques and our DL-based Symtosis^{TM} system show significant increase of liver segregation index for DL features compared to ML-based features. The system was validated using two class facial biometric data sets universally available.

Acknowledgement

The authors of National Institute of Technology Goa, India would like to acknowledge MediaLab Asia, Ministry of Electron-

ics and Information Technology, Government of India for their kind support.

Appendix A. Convolution Neural Network, DL and Inception Layer

A1. Convolution Neural Network

Convolution Neural Networks [16] have ability to decompose images into feature maps generating like a deck of cards representing the feature maps which can then be fed into limited layered neural networks for training. Mathematically, basic convolution can be represented as:

$$g(x,y) = I(x,y) \otimes w(x,y) = \sum_{s=-\frac{m}{2}}^{\frac{m}{2}} \sum_{t=-\frac{m}{2}}^{\frac{m}{2}} I(x+s, y+t) \quad \times w(x,y) \tag{A1}$$

Whereas image I is convolved with kernel w, yielding an output g, $\otimes$ represents the convolution operation. The convolution is basically sum of all products between image I and kernel w, represented by Eq. (A1), where the kernel is represented as a vector of size $m \times m$ and is shown for the point locations (x, y), while s and t are the dummy variables.

The ReLu is a simple operation where all negative pixel values in the feature map are set to zero using the maximization operation: $f(x) = \max(0, x)$, where, x is a unit in the feature map. It is like a comparator where the current value is being compared against 0 or a threshold value. This is also called ramp function and is analogous to half-wave rectification in electrical engineering. It is applied in CNNs to reduce the likelihood of the gradient to vanish. In our experiment, ReLu is implicitly combined with the convolution layer. The pooling reduces the dimensionality of each feature map but retaining the most important information i.e., max pooling and average pooling. Pooling is done to simplify the output from CNN. During dropout, reduction of over fitting takes place. At each training stage, individual nodes are either removed from the CNN with probability of "1-p" or kept with random probability p, so that a reduced network is left. The layers of DL architecture consisting of weights (parameters) to be trained are counted for depth calculation. Finally, the softmax loss is used for predicting a single class of L mutually exclusive classes:

$$\delta(z_j) = \frac{e^{z_j}}{\sum_{l=1}^{L} e^{z_l}} \text{ for } j = 1,\ 2,\ \ldots,\ L \tag{A2}$$

where, δ represents the loss and z is an L-dimensional class vector. It takes a vector of arbitrary real-valued scores and squashes it to a vector of values between zero and one that sums to one. In here, we see that class label of an instance is determined by maximum δ value among L number of δ values.

A2. DL architecture

The DL CNN architecture applied for US liver FLD detection and stratification is shown in Table 1. The most important part of this DL architecture is the application of inception layer that allows merging several convolutions and pooling operations into a single layer. Inception Model in the GPU framework was the sole contributor for higher performance. It does so by the application of dimensionality reduction using 1×1 convolutions. The inception layer allows the increasing of the depth of DL architecture without significant rise in computational cost.

The first convolution layer consists of 64 (7×7) filters, converting the image of size 229×229 image into $(112 \times 112) \times 64$ feature maps (SN 1). Max pooling is applied to these 64 feature maps to down sample them into $(56 \times 56) \times 64$ feature maps (SN 2). In here, 64 (1×1) convolutions are applied for dimensionality reduction (SN 3). A total of (3×3), 192 convolution filters are applied on the down sampled images to convert them into $(56 \times 56) \times 192$ feature maps (SN 3). Then max pooling is applied again to convert the feature maps into 192 (28×28) feature maps (SN 4). Two layers of inception, 3a and 3b converts these feature maps into, 480 (28×28) feature maps (SN 5 and SN 6). Max pooling is applied again to convert them into 480 (14×14) feature maps (SN 7). Now, five layers of inception, 4a, 4b, 4c, 4d and 4e are applied to convert these into 832 (14×14) feature maps (SN 8, 9, 10, 11 and 12, respectively). The images are down sampled using max pooling to convert them to 832 (7×7) feature maps (SN 13). Two layers of inception, namely 5a and 5b are applied to convert them into 1024 (7×7) feature maps (SN 14 and 15). Now, average pooling is applied to downs sample, these feature maps into 1024 (1×1) feature maps (SN 16). To prevent over-fitting, of the NN, 40% dropout is applied (SN 17). A linear layer is applied to adapt the feature maps into 1000 (1×1) label sets (SN 18). A softmax classifier layer of size 2 (1×1) yields us the final classification labels (SN 19). Thus nearly 1 million weights are computed during the cascaded layer.

A3. Inception layer

The concept of an inception layer is to prevent the complexity pile up due to larger sized convolutions filters on large sized US liver images. This is time intensive with a 22 layered model. This is time expensive and challenging to handle the bulk of data. It is therefore necessary to rely on 1×1 convolution filters prior to application of larger sized convolutions. This can be thought as a dimensionality reduction paradigm, i.e., changing the large size deck (say $28 \times 28 \times 192$) to smaller sized deck (say $28 \times 28 \times 16$) using 16 convolutions filters of size 1×1 (as shown in Table 1, SN 5). The second benefit of inception layer is to prepare itself for the subsequent neural network layer by presenting the feature maps at one culmination point. This can be achieved by running different sized convolution filters and then concatenating them (as seen in the Fig. 12). The concept of concatenation is shown by appending the four types feature maps, i.e., coming via the paths: (a), (b), (c) and (d) like a linked list as one stack, all available for the next layer to use. Note that feature maps via path (a), (b), (c) and (d) are already dimensionality reduced using 1×1 convolution filters prior to concatenation. Thus, both, (i) benefits of reduction and (ii) single culmination point is being leveraged in inception layer. The dimensionality reduction numbers can be more appreciated by checking Table 1, serial number (5) and (6).

Appendix B. Scientific validation of deep learning systems using biometric facial data sets

Scientific validation is always an integrated component of the system design. For validation, one needs to run another set of liver data sets whose results are known *a priori*. Since such a clinical data is hard to obtain, we use facial biometric data set to test the classification accuracy. Standardized well published facial database were adapted for our validation study. These databases are called Face94 [34] and Face95 [35] data bases. These databases have the two class biometric faces of males and female. Several publications are available in the literature who have demonstrated the usage of these two famous and standardized databases [39,40]. We therefore decided to use these two standardized databases.

We have conducted experiments to validate our results using Face94 and Face95 data set. The Face94 [34] data set consists 153 individual images with various expressions and poses seated at a

Table B1
Accuracy results using DL architecture on facial data sets.

SN	Data set	# of Training Samples	# of Testing Samples	# of Classes	# of Iterations	Avg. Accuracy (%)
1	Face94	620	160 (~20%)	2	36,000	99.4
2	Face95	1296	144 (~10%)	72	30,000	96.1

fixed distance from camera. There are 2 classes, male and female and total number of images are 780.

Standardized face 95 data set

The Face95 [35] data set consists 72 individual images with different expressions and poses seated at a fixed distance from camera. There are 72 classes both male and female and total number of images are 1440.

Classification accuracy on facial data sets

The results of the Deep Learning tissue characterization system when applied on Face94 and Face95 data set can be seen in Table B1. The accuracy results of the DL architecture using Face94 and Face95 data sets are: 99.4% and 96.1%, respectively. Note that the DL architecture was adapted in a multi-class framework. These accuracy numbers indicate our compatibility against the accuracy numbers when the DL architecture was tried on the liver data sets, demonstrating the same range reaching the optimal results. We further want to emphasize that the ratio of the training to testing sample (testing was nearly 20% in Face94 and 10% in Face95) was lower in facial data sets compared to liver data sets, which further validates the ability to classify accurately.

Appendix C. Feature extraction

C.1. *Haralick texture (GLCM)*

GLCM calculates the following features shown Table C1 from the co-occurrence matrix calculated **from the image.**

C.2. *Run length texture*

GRLM feature extraction algorithm calculate features from the run length matrix as shown in Table C2.

Table C2
Features from gray level run length matrix.

SN	Features	Description
1	Short Run Emphasis (SRE)	$\sum_{x=1}^{N_a}\sum_{y=1}^{N_b}\frac{R(x,y)}{y^2} / \sum_{x=1}^{N_a}\sum_{x=1}^{N_b}R(i,j)$
2	Long Run Emphasis (LRE)	$\sum_{x=1}^{N_a}\sum_{x=1}^{N_b}j^2R(x,y) / \sum_{x=1}^{N_a}\sum_{x=1}^{N_b}R(x,y)$
3	Gray level non uniformity	$\sum_{x=1}^{N_a}(\sum_{x=1}^{N_b}R(x,y)^2) / \sum_{x=1}^{N_a}\sum_{y=1}^{N_b}R(x,y)$
4	Run length Non-uniformity (RLNU)	$\sum_{x=1}^{N_a}(\sum_{x=1}^{N_b}R(x,y)^2) / \sum_{x=1}^{N_a}\sum_{x=1}^{N_b}R(x,y)$
5	Run Percentage (RP)	$\sum_{x=1}^{N_a}\sum_{x=1}^{N_b}\frac{R(x,y)}{S}$
6	Low Gray-level Run Emphasis (LGRE)	$\sum_{x=1}^{N_a}\sum_{y=1}^{N_b}\frac{R(x,y)}{x^2} / \sum_{x=1}^{N_a}\sum_{x=1}^{N_b}R(x,y)$
7	High Gray-level Run Emphasis (HGRE)	$\sum_{x=1}^{N_a}\sum_{x=1}^{N_b}x^2S(x,y) / \sum_{x=1}^{N_a}\sum_{x=1}^{N_b}S(x,y)$

C.3. *Gabor-based directional features*

Gabor filter is the combination of Gaussian and complex-plane wave edge detection filter. This combination tries to diminish the uncertainty in both spatial and frequency domains. The application of dilations and rotations of this function produce filters which helps in the alignment and scale-tunable edge and line detection. Gabor transform has an impulse response that can be represented by a sinusoidal wave (a plane wave for distinct frequency and aligned 2-D Gabor). The function is given as:

$$f(p,q) = \exp\left\{-\frac{1}{2}\left[\left(\frac{p}{\sigma_p}\right)^2 + \left(\frac{q}{\sigma_q}\right)^2\right]\right\}\exp[j2\pi(Vp + Wq)] \quad \text{(C1)}$$

Table C1
Features from gray level co-occurrence matrix.

SN	Features	Description	SN	Features	Description
1	Contrast	$\sum_{n=0}^{N-1} n^2 \sum_{i=0}^{N}\sum_{j=0}^{N} L(i,j)$	11	Sum average	$\sum_{i=2}^{2N} iL_{x+y}(i)$
2	Autocorrelation	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(ij)L(i,j)$	12	Sum entropy	$-\sum_{i=2}^{2N} iL_{x+y}(i)logL_{x+y}(i)$
3	Maximum probability	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} max(L(i,j))$	13	Sum variance	$\sum_{i=2}^{2N} i - I_{sent}{}^2 L_{x+y}(i)$
4	Dissimilarity	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\lvert i-j\rvert(L(i,j))$	14	Difference variance	$\sum_{i=2}^{2N} i - I_{savg}{}^2 L_{x-y}(i)$
5	Homogeneity	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} 1/(1+(i-j)^2)(L(i,j))$	15	Difference entropy	$-\sum_{i=2}^{2N} iL_{x-y}(i)logL_{x-y}(i)$
6	Entropy	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} L(i,j)logL(i,j)$	16	Information correlation measure 1	$HXY - HXY1\ /max(Hx - Hy)$
7	Energy	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} L(i,j)^2$	17	Information correlation measure 2	$1-(e^{-2HXY2\ -\ HXY})^{1/2}$
8	Correlation	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i,j)L(i,j)^2 - \mu_x\mu_y/\sigma_x\sigma_y$	18	Sum Of Squares Variance	$\sum_{x=1}^{m}\sum_{y=1}^{n} f(x,y)^2$
9	Cluster shade	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}((i+j)-\mu_x\mu_y)^3\ L(i,j)$	19	Inverse Difference	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} 1/(1+\lvert i-j\rvert(L(i,j))$
10	Variance	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i-\mu)^2 log\ (L(i,j)$			

where, (p, q) represents the spatial-domain rectilinear coordinates, (V, W) are the points that are the specific 2-D frequency of the complex sinusoid and (σ_p, σ_q) depict the spatial extent and bandwidth of f.

References

[1] M. Lazo, R. Hernaez, M.S. Eberhardt, S. Bonekamp, I. Kamel, E. Guallar, A. Koitesh, F.L. Brancati, J.M. Clark, Prevalence of nonalcoholic fatty liver disease in the United States: the third national health and nutrition examination survey, 1988–1994, Am. J. Epidemiol. 178 (2013) 38–45.

[2] J.D. Browning, L.S. Szczepaniak, R. Dobbins, J.D. Horton, J.C. Cohen, S.M. Grundy, H.H. Hobbs, Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity, Hepatology 40 (2004) 1387–1395.

[3] P. Lamb, D.V. Sahani, J.M. Fuentes-Orrego, M. Patino, A. Ghosh, P.R.S. Mendonça, Stratification of patients with liver fibrosis using dual-energy CT, IEEE Trans. Med. Imaging 34 (2015) 807–815.

[4] D. Guo, T. Qiu, J. Bian, W. Kang, L. Zhang, A computer-aided diagnostic system to discriminate SPIO-enhanced magnetic resonance hepatocellular carcinoma by a neural network classifier, Computer. Med. Imaging Graph. 33 (2009) 588–592.

[5] U.R. Acharya, J. Suri, Data mining framework for fatty liver disease classification in ultrasound: a hybrid feature extraction paradigm, Med. Phys. 39 (2012) 4255–4264.

[6] J.R. Quinlan, Simplifying decision trees, Int. J. Man-Mach. Stud. 27 (1987) 221–234.

[7] L. Saba, N. Dey, A.S. Ashour, S. Samanta, S.S. Nath, S. Chakraborty, J. Sanches, D. Kumar, R. Marinho, J.S. Suri, Automated stratification of liver disease in ultrasound: an online accurate feature classification paradigm, Comput. Methods Prog. Biomed. 130 (2016) 118–134.

[8] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Netw. 5 (1989) 359–366.

[9] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.

[10] S.C. Lo, H.P. Chan, J.S. Lin, H. Li, M.T. Freedman, S.K. Mun, Artificial convolution neural network for medical image pattern recognition, Neural Netw. 8 (1995) 1201–1214.

[11] V.N. Vapnik, An overview of statistical learning theory, IEEE Trans. Neural Netw. 10 (1999) 988–999.

[12] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (2006) 489–501.

[13] R. Ribeiro, R.T. Marinho, J. Velosa, F. Ramalho, J.M. Sanches, J.S. Suri, The usefulness of ultrasound in the classification of chronic liver disease, In Engineering in Medicine and Biology Society, in: EMBC, 2011 Annual International Conference of the IEEE, 2011, pp. 5132–5135.

[14] K. Gangadhar, K.N. Chintapalli, G. Cortez, S.V. Nair, MRI evaluation of fatty liver in day to day practice: Quantitative and qualitative methods, Egypt. J. Radiol. Nuclear Med. 45 (2014) 619–626.

[15] Y.M. Kadah, A.A. Farag, J.M. Zurada, A.M. Badawi, Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images, IEEE Trans. Med. Imaging. 15 (1996) 466–478.

[16] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444.

[17] N.M. Noor, J.C. Than, O.M. Rijal, R.M. Kassim, A. Yunus, A.A. Zeki, M. Anzidei, L. Saba, J.S. Suri, Automatic lung segmentation using control feedback system: morphology and texture paradigm, J. Med. Syst. 39 (2015) 22.

[18] P. McClure, A. Elnakib, M.A. El-Ghar, F. Khalifa, A. Soliman, T. El-Diasty, J.S. Suri, A. Elmaghraby, A. El-Baz, In-vitro and in-vivo diagnostic techniques for prostate cancer: A review, J. Biomed. Nanotechnol. 10 (2014) 2747–2777.

[19] U.R. Acharya, M.R. Mookiah, S.V. Sree, R. Yanti, R.J. Martis, L. Saba, F. Molinari, S. Guerriero, J.S. Suri, Evolutionary algorithm-based classifier parameter tuning for automatic ovarian cancer tissue characterization and classification, Ultraschall Med. 35 (2014) 237–245.

[20] L. Saba, P.K. Jain, H.S. Suri, N. Ikeda, T. Araki, B.K. Singh, A. Nicolaides, S. Shafique, A. Gupta, J.R. Laird, J.S. Suri, Plaque tissue morphology-based stroke risk stratification using carotid ultrasound: a polling-based PCA learning paradigm, J. Med. Syst. 41 (2017) 98.

[21] T. Araki, N. Ikeda, D. Shukla, P.K. Jain, N.D. Londhe, V.K. Shrivastava, S.K. Banchhor, L. Saba, A. Nicolaides, S. Shafique, J.R. Laird, J.S. Suri, PCA-based polling strategy in machine learning framework for coronary artery disease risk assessment in intravascular ultrasound: A link between carotid and coronary grayscale plaque morphology, Comput. Methods Programs Biomed. 128 (2016) 137–158.

[22] V. Kuppili, M. Biswas, A. Sreekumar, H.S. Suri, L. Saba, D.R. Edla, R.T. Marinhoe, J.M. Sanches, J.S. Suri, Extreme learning machine framework for risk stratification of fatty liver disease using ultrasound tissue characterization, J. Med. Syst. 41 (2017) 152.

[23] S.F. Huang, R.F. Chang, W.K. Moon, Y.H. Lee, D.R. Chen, J.S. Suri, Analysis of tumor vascularity using three-dimensional power Doppler ultrasound images, IEEE Trans. Med. Imaging 3 (2008) 320–330.

[24] B.K. Singh, K. Verma, A.S. Thoke, J.S. Suri, Risk stratification of 2D ultrasound-based breast lesions using hybrid feature selection in machine learning paradigm, Measur. J. (2017).

[25] G. Pareek, U.R. Acharya, S.V. Sree, G. Swapna, R. Yantri, R.J. Martis, L. Saba, G. Krishnamurthi, G. Mallarini, A. El-Baz, S. Al Ekish, M. Beland, J.S. Suri, Prostate tissue characterization/classification in 144 patient population using wavelet and higher order spectra features from transrectal ultrasound images, Technol. Cancer Res. Treat. 12 (2013) 545–557.

[26] V.K. Shrivastava, N.D. Londhe, R.S. Sonawane, J.S. Suri, A novel and robust bayesian approach for segmentation of psoriasis lesions and its risk stratification, Comput. Methods Prog. Biomed. (2017) (to appear).

[27] D.F. Specht, Probabilistic neural networks, Neural Netw. 3 (1990) 109–118.

[28] B. Carse, T.C. Fogarty, A. Munro, Evolving fuzzy rule based controllers using genetic algorithms, Fuzzy Sets Syst. 80 (1996) 273–293.

[29] L. Saba, M. Anzidei, C.N. de Cecco, M. Porcu, A. Balestrieri, R. Sanfilippo, M. Francone, A. Mereu, P. Lucatelli, R. Montisci, J.S. Suri, Carotid artery intra-plaque attenuation variability using computed tomography, Neurovascular Imaging 2 (2016) 12.

[30] S.K. Banchhor, N.D. Londhe, T. Araki, L. Saba, P. Radeva, J.R. Laird, J.S. Suri, Well-balanced system for coronary calcium detection and volume measurement in a low resolution intravascular ultrasound videos, Comput. Biol. Med. 84 (2017) 168–181.

[31] T. Araki, N. Ikeda, D. Shukla, P.K. Jain, N.D. Londhe, V.K. Shrivastava, S.K. Banchhor, L. Saba, A. Nicolaides, S. Shafique, J.R. Laird, PCA-based polling strategy in machine learning framework for coronary artery disease risk assessment in intravascular ultrasound: A link between carotid and coronary grayscale plaque morphology, Comput. Methods Prog. Biomed. 128 (2016) 137–158.

[32] V.K. Shrivastava, N.D. Londhe, R.S. Sonawane, J.S. Suri, Reliable and accurate psoriasis disease classification in dermatology images using comprehensive feature space in machine learning paradigm, Expert Syst. Appl. 42 (2015) 6184–6195.

[33] U.R. Acharya, S.V. Sree, M.M. Krishnan, N. Krishnananda, S. Ranjan, P. Umesh, J.S. Suri, Automated classification of patients with coronary artery disease using grayscale features from left ventricle echocardiographic images, Comput. Methods Prog. Biomed. 112 (2013) 624–632.

[34] Spacek, L., Face94 Database, URL: http://cswww.essex.ac.uk/mv/allfaces/faces94.html. (accessed 01.07.2017).

[35] Spacek, L., Face95 Database, URL: http://cswww.essex.ac.uk/mv/allfaces/faces95.html. (accessed 01.07.2017).

[36] U.R. Acharya, S.V. Sree, M.M. Krishnan, F. Molinari, W. ZieleŸnik, R.H. Bardales, A. Witkowska, J.S. Suri, Computer-aided diagnostic system for detection of hashimoto thyroiditis on ultrasound images from a polish population, Journal of Ultrasound Med. 33 (2014) 245–253.

[37] X. Liu, J.L. Song, S.H. Wang, J.W. Zhao, Y.Q. Chen, Learning to diagnose cirrhosis with liver capsule guided ultrasound image classification, Sensors 17 (2017) 149.

[38] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (2010) 1345–1359.

[39] P. Ying-Han, B.J. Teoh Andrew, C.L. Ngo David, A discriminant pseudo Zernike moments in face recognition, J. Res. Pract. Inf. Technol. 38 (2006) 197.

[40] N.G. Chitaliya, A.I. Trivedi, An efficient method for face feature extraction and recognition based on contourlet transform and principal component analysis using neural network, Int. J. Comput. Appl. (2010) 6.