GoogLeNet-Based Ensemble FCNet Classifier for Focal Liver Lesion Diagnosis

Lakshmipriya Balagourouchetty, *Member, IEEE*, Jayanthi K. Pragatheeswaran, Biju Pottakkat, and Ramkumar G

***Abstract*—Transfer learning techniques are recently preferred for the computer aided diagnosis (CAD) of variety of diseases, as it makes the classification feasible from limited training dataset. In this work, an ensemble FCNet classifier is proposed to classify hepatic lesions from the deep features extracted using GoogleNet–LReLU transfer learning approachs. In the existing GoogLeNet architecture three modifications are done: ReLU activation functions in the inception modules are replaced by leaky ReLU activation function; a stack of three fully connected layers are included before the classification layer; and deep features of different level of abstraction extracted from the output of every inception layer given as classifier input in order to significantly enhance the classifier performance. The performance of the proposed classifier by the virtue of the above mentioned modifications is tested on six classes of liver CT images namely normal, hepatocellular carcinoma, hemangioma, cyst, abscess and liver metastasis. The results presented in this work demonstrate the efficacy of the proposed classifier design in achieving better classification accuracy.**

***Index Terms*—Contrast Enhanced Computed Tomography, GoogLeNet, liver pathologies, multi-temporal fusion, transfer learning.**

I. Introduction

CONTRAST Enhanced Computed Tomography (CECT) has become an integral part of the therapeutic procedure of various liver pathologies including hepatocellular carcinoma (HCC), hemangioma, cirrhosis etc. A non-contrast liver CT image does not necessarily exhibit all significant variations between the normal and abnormal tissues of the liver. For the diagnosis of liver related diseases, abdominal CT acquisition is done thrice after the injection of contrast and hence the name triple phase CT. The three phases namely hepatic arterial (HA) phase, portal venous (PV) phase and delayed venous phase are acquired after a duration of 10 seconds, 30 to 40 seconds and 2 minutes respectively after the contrast injection [1]. The three phases of image carry distinct information about the pathology and doctors predominantly look in for the collective information contained in HA and PV phase images for liver disorder diagnosis.

Although the CECT images are traditionally used in identifying the liver pathologies, the diagnostic accuracy by visual interpretation is only suboptimal as in some cases the contrast differentiation between the healthy and lesion portions of the organ in the CT image is not too distinct and prominent. Also, occasionally the differential enhancement within the areas of tumour in CECT may sometimes get masked in the contrast, leading to ambiguous interpretation and this nudges a need for CAD of liver diseases. The CAD system is expected to serve two purposes: a) providing an enhanced version of CT image which distinctly highlights the presence of abnormal hepatic tissues among the normal ones; b) classifying the liver lesions conforming to the tissue type belonging to either normal or abnormal and if abnormal, the lesion type. As distinct information pertaining to liver pathology is present equally in both HA and PV phase images, it is very obvious that by combining them through multi-temporal fusion process, all relevant and significant information can be obtained from the single fused composite image. This is carried out in [2] by the authors of this paper as two stage enhancement process with multi-temporal fusion of HA and PV phase images as first stage followed by decorrelation stretching as second stage. A detailed explanation and the relevant outputs obtained out of this two stage enhancement process is presented in the previous work of this study by the authors Lakshmipriya *et al*. in [2] and is also briefed in Section III. As a continuation of [2], liver lesion categorization from the two stage enhanced output is carried out by the virtue of the classifier derived from the modified version of GoogLeNet architecture. Liver lesions generally fall under two categories viz. focal liver lesions, wherein the abnormal hepatic tissues will be present within a particular region in liver and diffused liver lesions in which the unhealthy tissues will be spread over the entire liver. Examples of diffused liver lesions are cirrhosis, fatty liver and lymphoma. The focal liver lesions are only considered for this study, out of which the HCC and metastasis are malignant and

Manuscript received May 30, 2019; revised August 28, 2019; accepted September 16, 2019. Date of publication September 20, 2019; date of current version June 5, 2020. This work was supported by the University Grants Commission, India, under Minor Research Project scheme under Grant F 6684/16. *(Corresponding author: Lakshmipriya Balagourouchetty.)*

L. Balagourouchetty is with the Department of Electronics and Communication Engineering, Manakula Vinayagar Institute of Technology, Puducherry 605 107, India (e-mail: lakshmipriyaece@pec.edu).

J. K. Pragatheeswaran is with the Department of Electronics and Communication Engineering, Pondicherry Engineering College, Puducherry 605014, India (e-mail: jayanthi@pec.edu).

B. Pottakkat is with the Department of Surgical Gastroenterology, Jawaharlal Institute of Postgraduate Medical Education & Research, Puducherry 605 006, India (e-mail: bijupottakat@gmail.com).

R. G is with the Department of Radio-Diagnosis, Jawaharlal Institute of Postgraduate Medical Education & Research, Puducherry 605 006, India (e-mail: gramk80@gmail.com).

Digital Object Identifier 10.1109/JBHI.2019.2942774

the other lesion categories like cyst, hemangioma and abscess are benign.

The rest of this paper is organized as follows: The literature survey is presented in Section II. Section III describes the system model and the associated works. Results and discussion are given in Section IV followed by concluding remarks in Section V.

II. Related Works

Image fusion, a process of integrating complementary information from two or more images into a single composite image is extensively utilized in the diagnosis of various types of cancer. This concept is applied in the enhancement stage of CT images in this paper. A transfer learning based modified classifier architecture is proposed in this work for the purpose of liver lesion classification. Customarily, classification of medical images is done using a traditional classifiers like artificial neural network [3], [4], support vector machine [5], [6], logistic regression [7] etc with hand-crafted features derived from the images. However, the classification accuracy in this classical method relies upon the type and quantity of features extracted and the complexity of the classifier design to suit the pattern recognition challenge. In recent times, deep learning has become a thriving force in the classification of medical images, as it solves increasingly complicated challenges with increasing accuracy over time [8]–[10].

Deep convolutional neural networks (CNNs) avoids the extraction of handcrafted features and learns features at different levels very deeply by itself and therefore have manifested themselves as an effective tool for the classification of liver related disorders [11]–[13]. In [11], authors have implemented CNN based deep learning architecture to effectively classify cysts, metastases and hemangioma. Due to the limited availability of dataset for liver lesions, the authors have synthesized high quality liver lesions using generative adversarial networks (GAN) and showed remarkable improvement in the performance [11]. Enhanced detection capabilities for the detection of liver metastasis is achieved in [12] using fully convolution network (FCN) based on patch level analysis with superpixel sparse based classification. Watershed Gaussian based deep learning technique is proposed in [13] by combining the traditional feature extraction technique and deep neural network to classify three different classes of liver cancer namely HCC, hemangioma and metastasis and achieved good accuracy. Though, deep learning has recorded good performance in classification of variety of images, the training dataset requirement is huge to obtain a high degree of accuracy. This serves as the limitation for being popularly used in medical field [14].

Transfer learning is another paradigm shift in technology which overcomes the above mentioned challenge by using a pretrained deep neural network to train and classify any image datasets. A diverse set of such pretrained networks namely Alexnet [15], GoogLeNet [16], VGGNet [17] & [18], CifarNet [19], AggNet [20] and so on are developed to compete successfully in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Later these networks were used in the medical image recognition research [21]–[24] in the literature and accomplished phenomenal results. A deeper version of GoogLeNet – CNN has been fine-tuned and trained on thoracoabdominal lymph node datasets and interstitial lung disease dataset in [21]. In [22], the most progressive transfer learning architecture GoogLeNet has been fine tuned to classify lung cancer images. Liver fibrosis classification using transfer learning approach using VGGNet and fully connected neural network (FCNet) for ultrasound images was done in [23] with an accuracy of 96.06% due to the presence of three fully connected layers used for the classification. Deep learning approach using the existing architecture of GoogLeNet was implemented by the authors in [20] for the classification of HCC and normal liver samples extracted from diagnostic information bearing phases viz. HA and PV phase individually with an accuracy of 92.08%. In this work, the HA and PV phase images are considered as separate entity for training and testing the classifier. Inspired by the performance of various deep learning architectures in medical image recognition tasks, in this paper ensembling based modified GoogLeNet framework for liver lesion diagnosis is proposed. The two stage enhanced outputs [2] which has proven itself to be effective in the visual interpretation during diagnosis are applied as input to the classifier. To quantify the significance of two stage enhancement technique, the classifier performance is also evaluated with HA and PV phase images considered independently as input. The main contributions of this work towards the improvement in classification accuracy are as follows:

a) Complete diagnostic information present in both HA and PV phase images are collectively extracted from the two stage enhanced image using GoogLeNet based deep neural network.
b) Deep feature extraction from all nine different inception layers of the network both individually and aggregation of these features by means of our proposed ensemble FCNet classifier.
c) The existing GoogLeNet architecture is modified by altering the activation function in the inception layers of the network, thereby achieving a classification accuracy as high as 97.37%.
d) The proposed framework is evaluated on six classes of liver images comprising of normal liver and five focal liver lesions unlike the works in literature which have restricted themselves to a maximum of three classes [11], [23] & [24].

III. Proposed CAD Framework

The block diagram of the proposed CAD framework is shown in Fig. 1. This system consists of three modules: liver CT image enhancement, segmentation and liver lesion classification.

In the first module, enhancement is achieved in two stages: first, by means of multi-temporal image fusion of HA and PV phase CT and the second stage by using the decorrelation stretching operation. This is clearly elucidated in [2]. In the second module, liver lesions are segmented by means of dynamic colour gradient thresholding algorithm, followed by lesion classification by means of transfer learning through a modified GoogLeNet architecture discussed in III D.

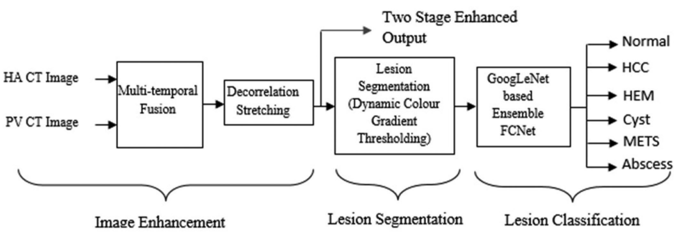

Fig. 1. Proposed CAD framework for liver lesion diagnosis.

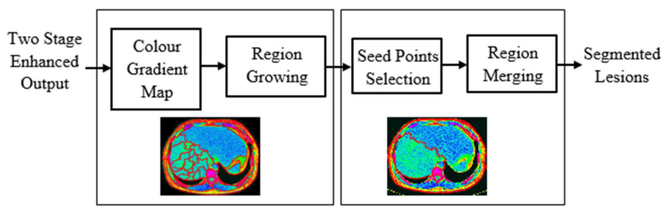

Fig. 2. Modules of lesion segmentation with output images.

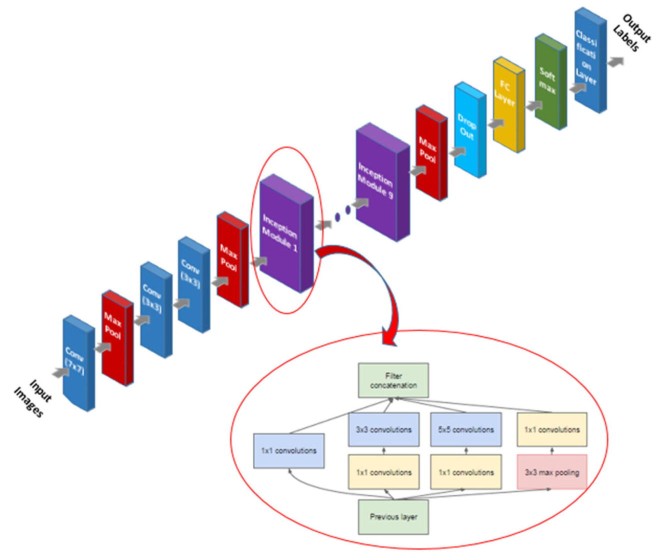

Fig. 3. GoogLeNet architecture.

A. Lesion Segmentation

Prior to classification, regions of interest corresponding to lesions have to be extracted from the two stage enhanced colour fused CT image. Region based segmentation using dynamic colour gradient thresholding presented in [25] is adopted in this work to extract the liver lesions. This segmentation algorithm is carried out in two steps namely region growing and region merging and the corresponding stepwise outputs are shown in Fig. 2. The region growing phase of the segmentation algorithm groups the regions with similar pixels and to begin this process, seed points are required. To automatically select the seed points, a colour gradient map D, which is a matrix of gradient magnitude at each pixel location of the input coloured image is formulated using (1).

$$D = \begin{bmatrix} du/dx & du/dy \\ dv/dx & dv/dy \\ dw/dx & dw/dy \end{bmatrix} \tag{1}$$

where, u, v and w are the three channels of the two stage enhanced colour fused CT image.

From the colour gradient map, the pixels in the homogeneous regions are selected as seeds for region growing and are selected using Otsu's thresholding technique [26]. The dynamic thresholding algorithm [25] ascertains that the grown regions are void of strong edges. The region growing segmentation process continues until there is no pixel left unsegmented. Since the seed regions are automatically selected based on colour gradients, there is a possibility of over segmentation i.e., similar regions present at different locations within the image are grouped as different partitions. Also in most cases of liver pathology except for cyst and hemangioma, the lesions will be scattered and exhibit some sort of discontinuity within the liver and this would have been grouped as different regions in the previous region growing algorithm. Hence, there arises a need to merge the disjoint lesions masking all other regions of the image. Therefore, in this work, the seed points belonging to the liver lesions are manually selected by the medical experts, who are the authors of this work and merged. Thus the manual seed point selection belonging to the lesion regions also ensures that no lesion is left during the segmentation process.

B. GoogLeNet

GoogLeNet [16] is a popular pre-trained deep convolutional neural network (CNN) architecture with 22 layers developed to outperform the existing CNN architectures. It manifested itself to be superior in perceiving the visual patterns precisely from the source and won ILSVRC challenge in the year 2014.The credit for the tremendous performance improvement goes to the inception modules (IM) (highlighted by red circle in Fig. 3) embodied within its architecture [16].

IM which permits multiple convolutions with multiple kernels and max pooling to take place simultaneously within a single layer ensures that the network trains with optimal weights and selects more useful features. To do this, every inception layer encompasses within itself, variable size convolutional kernels namely 1×1, 3×3 and 5×5 and an additional 3×3 max pooling to capture more discriminative features from the pattern passed from previous layer. The architecture of IM is also shown in Fig. 3 wherein the 1×1 filters along with max pooling layer performs dual task of dimensionality reduction and summarizing the contents from previous layer. As shown in Fig. 3, Nine IMs in the GoogLeNet architecture effectively extract the more discriminative features from the source image. However IMs of GoogLeNet architectures possess certain drawbacks and unexplained complexity [27]. Generally features with higher dimensional representations are preferred as they will be easy to handle in a new network and contribute more towards speedy training process [27]. This is contradictory to the role of IM of GoogLeNet architecture which selects more optimized features with reduced dimensions. To overcome this shortfall, in [27] larger convolutional filters are replaced by mini – networks implemented by factorizing the larger convolutional kernels in smaller ones. In other words, any n $\times$ n convolution can be implemented by 1 $\times$ n convolution succeeded by n $\times$ 1 convolution. Though this factorization operation theoretically holds good for all values of n, in practice it works good only for middle layers and not for initial inception layers. Additionally, like any other deep neural network, GoogLeNet architecture as well suffers from vanishing gradient problem [27]. This work attempts to tackle both dimensionality reduction and vanishing gradient problem at a stretch with suitable modifications which is claimed as the significant contribution. In order to address the issue of dimensionality reduction through inception layers, the features from each of the inception layers are aggregated prior to classification. Furthermore to alleviate the vanishing gradient problem, the ReLU activation function in the IM is replaced by leaky ReLu (LReLu) activation function explained in the subsequent section. Incorporating these modifications, in this work a new classifier is derived from the existing GoogLeNet architecture and is named as Ensemble FCNet classifier. The corresponding modification done in the existing GoogLeNet architecture is clearly presented in Section III D.

C. GoogLeNet – LReLU

Activation functions play a very crucial role of deciding decides the firing of neuron in network learning A non-linear activation function, rectified linear unit (ReLU) given in Equation (2) is computationally simple and has demonstrated to be effective especially in deep neural networks.

$$g\left(z\right) = \max\left\{0, z\right\} \tag{2}$$

The neuron firing condition is given in (3)

$$h = g\left(W^T x + b\right) \tag{3}$$

where h is the output of the neuron, W is the weight matrix associated with the inputs x to the neurons and b is the bias. However Equation (2) gives no assurance that all the neurons will be active throughout, as they do not learn and update via the optimization algorithm and fail to fire the neuron with negative inputs thereby making all the neurons following it to be dead. This problem is called dying ReLU, which is undesirable as it progressively makes a major portion of the network to be inactive. This issue can be overcome by assigning a small positive value such as 0.1 instead of 0 in Equation (2) to make all the neurons to be active for most of the training samples and allow the inputs to pass through and thereby solving the vanishing gradient problem. This generalization in ReLU function is given below in Equation (4).

$$h_i = g(z,a)_i = \max\left(0, z_i\right) + a_i \min\left(0, z_i\right) \tag{4}$$

where a_i a non-zero slope called leak factor is assigned to the non-linear activation function. The value of a_i is zero for ReLU. This generalization is called leaky ReLU (LReLU) and the typical value of a_i is 0.01. The transfer function of ReLU and LReLU activation functions are shown in Fig. 4 for comparison purpose.

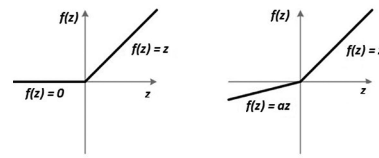

Fig. 4. Transfer function of ReLU and leaky ReLU activation functions.

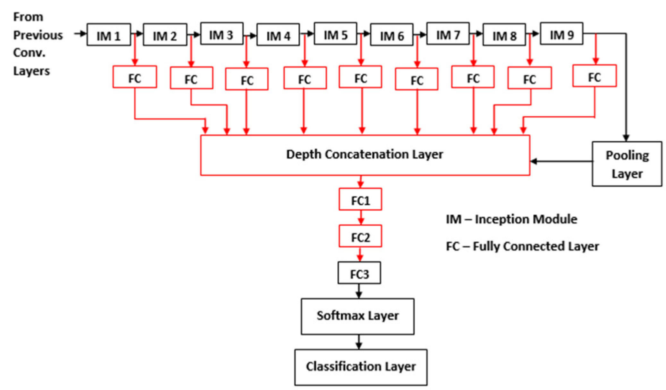

Fig. 5. Architecture of the proposed ensemble FCNet classifier design.

Mathematically, the leaky ReLU (LReLU) activation function is given as

$$\text{when } z_i < 0 : h_i = g(z,a)_i = a_i.z_i \tag{5}$$

The leak helps in increasing the range of the ReLU function, thereby ensuring every neuron of the deep network to contribute for the active performance of the deep network. Since the performance of GoogLeNet is claimed to be greatly influenced by the IM, all the ReLU activation functions within the IMs are replaced by LReLU [28] in this work. This modification is done using deepNetworkDesigner app present in deep learning toolbox of MATLAB 2018b.

D. Proposed Classifier

In this work, a new classifier derived from GoogLeNet called ensemble fully connected network (FCNet) as shown in Fig. 5 with the modifications done indicated in red colour is proposed to improve the classification accuracy (reason for its name is justified at the end of this section).

It has been stated already in Section III.B that every IM present in GoogLeNet architecture by virtue of multiple parallel connected convolution filters and pooling layers, extracts deep high level features pertaining to the input image. Now that with nine different IMs extracting features at different levels of abstraction, these features can be effectively combined before classification by a fully connected layer to achieve higher degree of classification accuracy. Since the dimensions of feature vector vary for every IM, a separate fully connected layer is used at the output of every IM in this work and finally the outputs of all the fully connected layers are concatenated using a depth concatenation layer. In a conventional deep neural network, series of convolutional layers and maxpooling layers perform

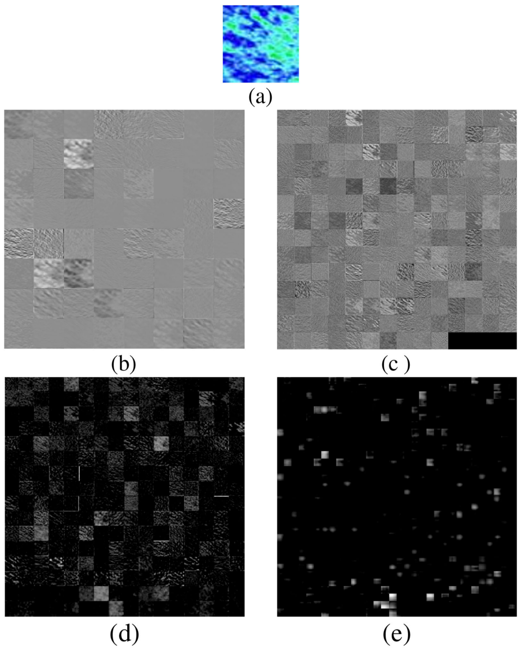

Fig. 6. Illustration of low and high level features extracted. (a) Input sample (b) Outputs of 1st layer (convolution 7 × 7) (c) Low level features extracted from 1st Inception Layer (d) Mid – level features extracted from 4th inception layer (e) High level features extracted from 9th inception layer.

feature extraction and a fully connected layer uses the extracted features for the task of classification. Fully connected layer is the traditional multilayer perceptron neural network with softmax activation function. The modified architecture of GoogLeNet extracts deep features from the input images and are then classified using a deep neural network comprised of three fully connected layers referred to as FCNets [23]. The numbers of FCNets used for classification is restricted to 3 to prevent overfitting. Fig. 6 shows the montage view of low, mid and high level features extracted from the 1st, 4th and 9th inception layers for one sample of HCC lesion. This gives a powerful insight about features with various levels of abstraction and justifies the need for feature ensembling proposed which is the major contribution of this work.

Due to the fact that the deep features extracted individually from nine inception layers are effectively combined through fully connected layers and given as input to the stack of FCNets, the proposed classifier is named as ensemble FCNet classifier. Ensembling is a technique used to machine learning to combine several entities (like feature, classifier etc), to produce a single optimal entity. The ensemble methodology used in this work is simple aggregation ensemble method which combines the features from the output of every inception layer.

LReLU is chosen to be the activation function for all the fully connected layers except the last one in the proposed classifier. The last fully connected layer in the proposed classifier design uses softmax activation function to perform the final classification by the virtue of probability distribution classification for each target class. The performance of the proposed classifier outperforms the state–of–the–art transfer learning architectures namely GoogLeNet and Alexnet, which is evident from the discussion on the obtained results.

TABLE I
DATA SET CATEGORY WITH COUNT

Class		No. of images
Healthy Liver	Normal	105
Malignant Tumour	HCC	134
	METS	77
Benign Tumour	HEM	110
	Cyst	103
	Abscess	105
Total		**634**

IV. Results and Discussion

A. Dataset

The images used in this study are obtained from Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry, India with ethical clearance. CT scans were acquired using Siemens Sensation 64 detector scanner via a standard four phase contrast enhanced imaging protocol. The obtained images are in DICOM (Digital Imaging and Communications in Medicine) format and the required slices are converted to JPG format using RadiAnt DICOM viewer software [29]. All the images considered are of uniform size 512 × 512 pixels. Totally 634 liver CT images with 70% and 30% for training and testing purposes as mentioned in Table I are considered for this work. The images used for the purpose of enhancement and classification in this work fall under three major categories: a) normal liver i.e., healthy liver, b) malignant lesions comprising of HCC and METS and c) benign lesions namely HEM, cyst and abscess. The classifier performance is assessed for four different set of input images viz. raw HA and PV phase images, multi-temporal fused images and two stage enhanced output images (Refer Fig. 7).

B. Review on Two-Stage Enhanced Outputs

Although the two–stage enhancement module is the previous work of the authors, the enhanced images are presented here again for ready reference along with conventional HA and PV phase images as they are fed as the inputs for the classification module.

It is seen in Fig. 7, the decorrelation stretching process of the enhancement module assigns different colours for different portions in the liver CT image. Normal liver is assigned blue colour, the blood vessels within the liver are assigned pink colour and the lesions are assigned green colour with varying hues and intensity. This manifests the ability of the proposed enhancement approach in assisting the surgeons in their diagnostic procedure with a high degree of discrimination between the healthy and unhealthy liver tissues. The enhanced image by the virtue of its colour has a significant impact in achieving better segmentation accuracy and also in classification accuracy.

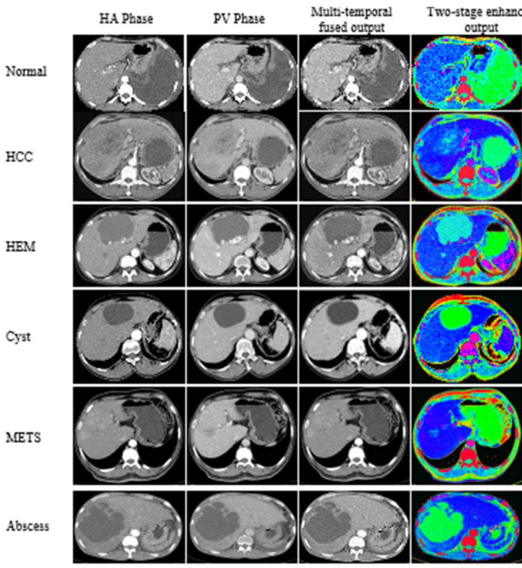

Fig. 7. Two stage enhanced output images.

TABLE II
TRAINING PARAMETERS

Sl No	Parameter	Value
1	No. of Epochs	40
2	Learning Rate	0.0001
3	Hyperparameter Tuning Algorithm	Adam

C. Assessment of Proposed Classifier

The improvement in classification accuracy owing to the two stage enhancement module and modifications done in the GoogLeNet architecture are presented in this section. Most of the deep learning architecture in literature for liver lesion classification have been restricted to a maximum of three classes of lesions [11], [23] & [24]. Against this, in this work five lesion classes along with the normal liver are considered for the purpose of classification. The training parameters for classification using GoogLeNet transfer learning technique are tabulated in Table II.

GoogLeNet, being a pre-trained network, the architecture comes with the inherent pre-trained weights and during the process of fine tuning the network to liver dataset used in this work, the weights are updated using adaptive learning rate optimization algorithm [30] popularly called as 'adam' with an initial learning rate of 0.0001. Training progress for the case of proposed classifier and that of GoogLeNet is shown in Fig. 8(a) & (b).

It is seen from the Fig. 8, that the training accuracy for the case of proposed classifier converges at seventeenth epoch, unlike the GoogLeNet which converges at thirty second epoch. This elucidates the ability of the proposed classifier to learn from deep features quickly owing to higher dimensional representation as mentioned in Section III B. The comparison of the classifier performance for the enhanced and raw liver CT images are tabulated in Table III.

In addition, the performance of the proposed classifier is compared with that of Alexnet, GoogLeNet transfer learning techniques and GoogLeNet with SVM classifier for all four sets of input images as mentioned above and the comparison of classification accuracies are presented in Table IV. Additionally, the enhanced performance of the proposed classifier is demonstrated graphically by means of receiver operating characteristics (ROC) plot shown in Fig. 9, which is a plot of true positive rate (TPR) Vs false positive rate (FPR). It is observed that the area under the curve (AUC) in Fig. 9 is highest for the proposed ensemble classifier, and hence proves to be efficient.

The modification of GoogLeNet architecture implemented in this work, GoogLeNet – LReLU inception layers along with the ensemble FCNet classifier has shown a remarkable improvement in classifying the liver lesions for all six dataset of liver images. From the metrics presented in Tables III and IV, it is evident that in addition to classification, the classifier plays another crucial role of quantifying the breadth of enhancement combinedly achieved by means of multi-temporal fusion and decorrelation stretching, which significantly contributes towards achieving better accuracy. The boost in classification accuracy of the proposed classifier and the other two transfer learning methods (refer Table IV) is visible when the two stage enhanced output images are given to the classifier. In all the cases, it is seen that the two stage enhancement module has a greater impact in achieving better classification accuracy for all classifier types considered in this study. The overall accuracy of the proposed classifier model derived from the two stage enhanced image records a far higher value in comparison with multi-temporal fused images and HA phase images. The improvement in the proposed classifier performance is proclaimed to be achieved due to the following reasons: 1) by virtue of decorrelation stretching and fusion of two informative HA & PV phases of images and 2) due to the deep features extracted from every IMs of modified GoogLeNet architecure. To investigate the impact of features at the output of various IMs, the classifier performance with different combinations of features from various IMs are assessed with LReLU activation functions for the two stage enhanced images category and is presented in Table V. It is evident from the results in Table V that the classifier performance significantly improves when low, medium and high level features are combined and is the highest when all IM outputs are combined. This justifies the ensembling of features from all 9 IMs considered in the proposed modified GoogLeNet architecture.

The discussion presented so far on the proposed classifier elucidates its supremacy when compared with other existing classifier options. In addition to this, it is also essential to compare the obtained results with the hepatic lesion classification metrics reported in the literature [11], [23] & [24]. There exists no standard database available for liver lesion classification and most of the work are implemented on the images obtained by the corresponding authors from hospitals present in their locality and therefore working on the same datasets of the other earlier published works [11], [23] & [24] is not possible. Accordingly, the summary of the results of recent research on liver ailment classification using deep learning techniques along with the results obtained in this study are presented in Table VI for comparison purpose.

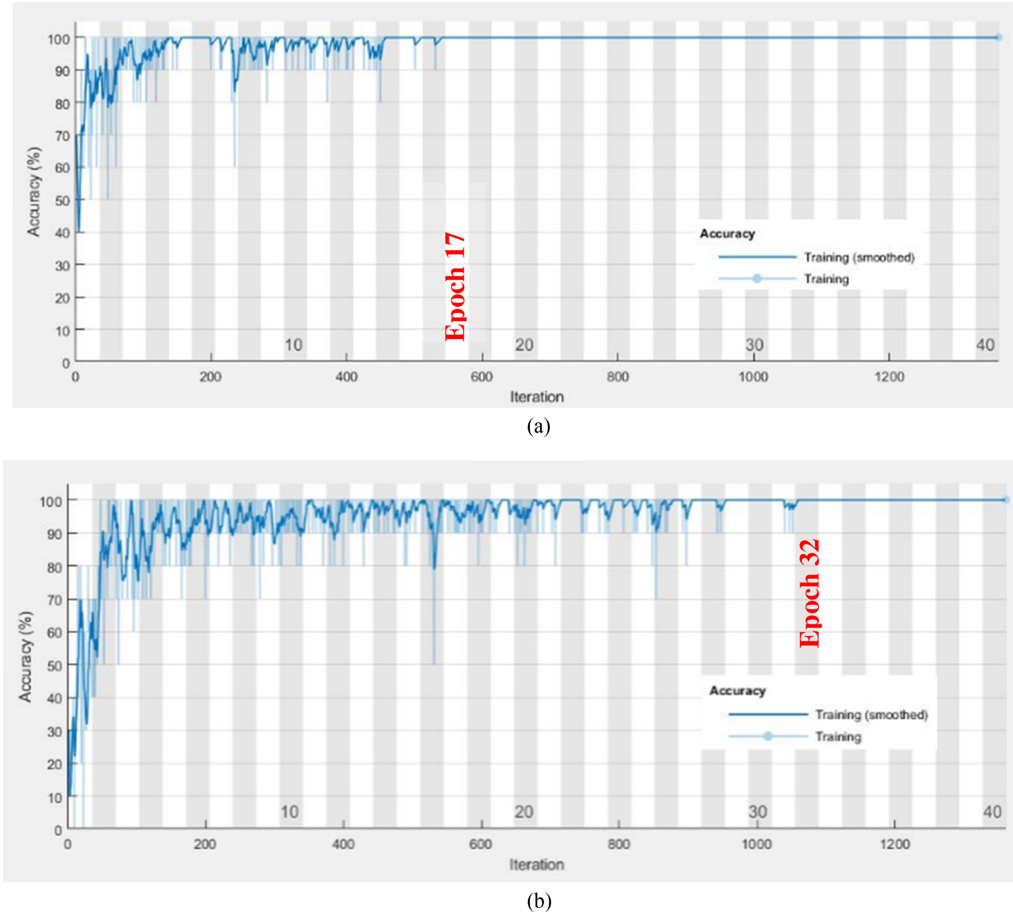

Fig. 8. Training progress of (a) proposed classifier and (b) GoogLeNet.

TABLE III
PROPOSED CLASSIFIER PERFORMANCE

Input Images	Classification Accuracy
Two Stage Enhanced Images	**97.37%**
Multi-temporal Fused Images	93.75%
HA Phase Images	89.58%
PV Phase Images	87.22%

TABLE IV
COMPARISON OF PROPOSED CLASSIFIER WITH EXISTING TRANSFER LEARNING APPROACHES

Classifier Types	HA Phase Images	PV Phase Images	Multi-temporal Fused Images	Two Stage Enhanced Images
Proposed Classifier	**89.58%**	**87.22%**	**93.75%**	**97.37%**
GoogLeNet – LReLU	88.01%	85.48%	92.53%	96.26%
GoogLeNet	83.22%	82.01%	89.95%	93.63%
GoogLeNet + SVM	82.62%	80.79%	86.59%	89.99%
Alexnet	80.55%	77.34%	82.67%	86.66%

Two important inferences can be derived from Table VI: i) The maximum number of lesion classes considered in the literature is restricted to three unlike our case it is five lesions and one normal category, ii) The obtained performance metric is also the best among all, in spite of the increased number of lesion categories. These are the strengths of this research study.

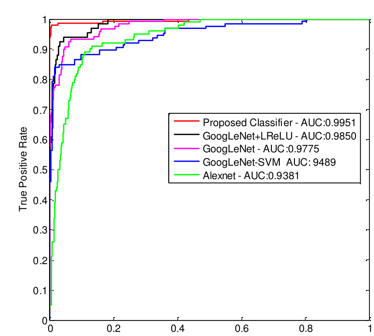

Fig. 9. ROC plot depicting classifier performance.

D. Research Findings

This section is intended towards presenting the major highlights of this work on liver lesion classification. To summarize, the following are the observations derived from the obtained results:

TABLE V
CLASSIFIER PERFORMANCE FOR DIFFERENT LEVELS OF FEATURES

Inception Module Combination	Feature Level	Classification Accuracy
$IM_1 – IM_9$ (Ensemble FCNet Classifier)	**Low, Medium, High**	**97.37%**
IM_1,IM_3,IM_5,IM_7,IM_9	Low, Medium, High	97.02%
IM_6,IM_7,IM_8, IM_9	Medium, High	96.72%
IM_1, IM_5, IM_9	Low, Medium, High	96.57%
IM_1, IM_2, IM_9	Low, High	96.34%
IM_9 (GoogLeNet LReLU)	High	96.26%

TABLE VI
SUMMARY OF PERFORMANCE METRICS FOR HEPATIC LESION CLASSIFICATION IN LITERATURE

Reference	Lesion Class	Deep Learning Network	Performance	
			Metric	Value
[11]	Cyst, METS HEM	CNN - GAN	Sensitivity Specificity	85.7% 92.4%
[23]	Liver Fibrosis	VGGNet + FCNet	Accuracy	96.06%
[24]	Normal HCC	GoogLeNet	Accuracy	92.8%
Proposed	*Normal & 5 Lesions*	GoogLeNet	Accuracy	93.63%
	Normal & 5 Lesions	*GoogLeNet - LReLU*	*Accuracy*	*96.26%*
		GoogLeNet – LReLU + Ensemble FCNet Classifier	*Accuracy*	*97.37%*

i) The two stage enhancement modulde in addition to providing assistance in precise visual diagnosis also influences the classification performance significantly.
ii) The modification in terms of activation function of all nine IMs alone has resulted in a good classification accuracy of 96.26%.
iii) The proposed classifier by the virtue of deep features extracted at various level of abstraction along with LReLu activation function contributes significantly in achieving a classification accuracy of 97.37% for five lesion classes along with the normal liver category.

V. CONCLUSION

In this paper, a novel framework for the liver lesion classification using GoogLeNet based ensemble FCNet classifier is proposed. The classification is performed on the enhanced images, which are subjected to a two stage enhancement process that involves two operations: multi-temporal fusion and decorrelation stretching. Deep aggregate features extracted from the two stage enhanced image at nine different intermediate levels of GoogLeNet architecture very well represent the inherent characteristics of the different types of lesions. Further, the modification in the GoogLeNet inception layer architecture through LReLU activation function also assists in identifying the discriminating components among the lesion types considered herewith. These features combined with the proposed classifier records a prominent performance in classifying the 5 classes of liver lesions along with healthy liver. The proposed classifier outperforms the existing state-of-the art classifiers and reports 4% improvement compared with the most acclaimed GoogLeNet transfer learning approach. Besides classification, the classifier acts as an additional tool of quantifying the effectiveness of the two stage enhancement module. The results presented in Section IV, epitomize the proficiency of this framework in providing a precise visual information about the presence of tumourous hepatic tissues and achieving a greater degree of accuracy in hepatic lesion classification. Henceforth, the proposed diagnostic framework is claimed to offer an elite assistance for the medical experts by offering timely and precise diagnostic impression.

ACKNOWLEDGMENT

The authors would like to thank Dr. K. Sasikumar and Dr. R. Gurram, Senior Residents, Department of Surgical Gastroenterology, Jawaharlal Institute of Postgraduate Medical Education & Research, Puducherry, India, for extending their support toward obtaining the CT images used in this study.

REFERENCES

[1] S. Roy, Y. Chi, J. Liu, S. K. Venkatesh, and M. S. Brown, "Three-dimensional spatiotemporal features for fast content-based retrieval of focal liver lesions," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 11, pp. 2768–2778, Nov. 2014.

[2] B. L. Priya, S. J. Adaikalamarie, and K. Jayanthi, "An effective liver cancer diagnosis through multi-temporal fusion and decorrelation stretching techniques," in *Proc. Int. Conf. Comput. Intell. Commun.*, Oct. 2016, vol. 14, pp. 22–26.

[3] E. L. Chen, P. C. Chung, C. L. Chen, H.-M. Tsai, and C.-I. Chang, "An automatic diagnostic system for CT liver image classification," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 6, pp. 783–794, Jun. 1998.

[4] M. Gletsos, S. G. Mougiakakou, G. K. Matsopoulos, K. S. Nikita, A. S. Nikita, and D. Kelekis, "A computer aided diagnostic system to characterize CT focal liver lesions: Design and optimization of a neural network classifier," *IEEE Trans. Inf. Technol. Biomed.*, vol. 7, no. 3, pp. 153–162, Sep. 2003.

[5] I. Diamant J. Goldberger, E. Klang, M. Amitai, and H. Greenspan, "Multi-phase liver lesions classification using relevant visual words based on mutual information," in *Proc. Int. Symp. Biomed. Imag.*, Jul. 2015, pp. 407–410.

[6] Y. L. Huang, J. H. Chen, and W. C. Shen, "Diagnosis of hepatic tumors with texture analysis in nonenhanced computed tomography images," *Acad. Radiol.*, vol. 13, no. 6, pp. 713–720, Jun. 2006.

[7] C. C. Chang *et al.*, "Computer-Aided diagnosis of liver tumors on computed tomography images," *Comput. Methods Programs Biomed.*, vol. 145, pp. 45–51, Jul. 2017.

[8] S. Duraisamy and S. Emperumal, "Computer-aided mammogram diagnosis system using deep learning convolutional fully complex-valued relaxation neural network classifier," *IET Comput. Vision*, vol. 11, no. 8, pp. 656–662, Dec. 2017.

[9] E. Trivizakis *et al.*, "Extending 2-D convolutional neural networks to 3-D for advancing deep learning cancer classification with application to MRI liver tumor differentiation," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 3, pp. 923–930, May 2019.

[10] L. Yuan, X. Wei, H. Shen, L. Zeng, and D. Hu, "Multi-center brain imaging classification using a novel 3D CNN approach," *IEEE Access*, vol. 6, pp. 49925–49934, 2018.

[11] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, Dec. 2018.

[12] A. Ben-Cohen *et al.*, "Fully convolutional network and Sparsity-Based dictionary learning for liver lesion detection in CT examinations," *Neurocomputing*, vol. 275, pp. 1585–1594, Jan. 2018.

[13] A. Das, U. Rajendra Acharya, S. S. Panda, and S. Sabut, “Deep learning based liver cancer detection using watershed transform and gaussian mixture model techniques,” *Cogn. Syst. Res.*, vol. 45, pp. 165–175, May 2019.

[14] A. Lumini and L. Nanni, “Deep learning and transfer learning features for plankton classification,” *Ecological Inform.*, vol. 51, pp. 33–43, May 2019.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2012, pp. 1097–1105.

[16] C. Szegedy *et al.*, “Going deeper with convolutions,” in *Proc IEEE Conf. Comput. Vision Pattern Recognit.*, 2015, pp. 1–9.

[17] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, “Overfeat: Integrated recognition, localization and detection using convolutional networks,” in *Proc. Int. Conf. Learn. Representations (ICLR)*, Dec. 2013.

[18] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *Proc. Int. Conf. Learn. Representations (ICLR)*, Apr. 2015.

[19] A. Krizhevsky, “Learning multiple layers of features from tiny images,” Master thesis, Dept. Comput. Sci., Univ. Toronto, Toronto, ON, USA, 2009.

[20] S. Albarqouni, C. Baur, F. Achilles, V. Belagiannis, S. Demirci, and N. Navab, “AggNet: Deep learning from crowds for mitosis detection in breast cancer histology images,” *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1313–1321, May 2016.

[21] H. Shin *et al.*, “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” *IEEE Trans. Medical Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.

[22] T. Fang, “A novel computer-aided lung cancer detection method based on transfer learning from GoogLeNet and median intensity projections,” in *Proc. IEEE Int. Conf. Comput. Commun. Eng. Technol.*, Aug. 2018, pp. 286–290.

[23] D. Meng, L. Zhang, G. Cao, W. Cao, G. Zhang, and B. Hu, “Liver fibrosis classification based on transfer learning and FCNet for ultrasound images,” *IEEE Access*, vol. 5, pp. 5804–5810, Mar. 2017.

[24] A. Brunetti, L. Carnimeo, G. F. Trotta, and V. Bevilacqua, “Computer-assisted frameworks for classification of liver, breast and blood neoplasias via neural networks: A survey based on medical images,” *Neurocomputing*, vol. 335, pp. 274–298, Mar. 2019.

[25] G. Balasubramani *et al.*, “Unsupervised color image segmentation using a dynamic color gradient thresholding algorithm,” *Proc. SPIE*, vol. 6806, Feb. 2008, Art. no. 68061H.

[26] N. Otsu, “A threshold selection method from gray-level histogram,” *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.

[27] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2016, pp. 2818–2826.

[28] D. Kingma and J. Ba, “A method for stochastic optimization,” in *Proc. Int. Conf. Learn. Representations*, Jan. 2017.

[29] [Online]. Available: https://www.radiantviewer.com/products/radiant-dicom-viewer-standard/, Accessed on: Dec. 2018.

[30] A. L. Maas, A. Y. Hannun, and A. Y. Ng, “Rectifier nonlinearities improve neural network acoustic models,” in *Proc. 30th Int. Conf. Mach. Learn.*, 2013, pp. 1–6.