Multimedia Tools and Applications (2022) 81:1607–1623
https://doi.org/10.1007/s11042-021-11411-z

FireNet-MLstm for classifying liver lesions by using deep features in CT images

Gedeon Kashala Kabe[1] · **Yuqing Song**[1] · **Zhe Liu**[1]

Received: 5 October 2020 / Revised: 16 January 2021 / Accepted: 2 August 2021 /
Published online: 8 October 2021

Abstract
Nowadays working in the medical imaging domain remains a big challenge, since the collecting such datasets is so complex, deep learning techniques have accomplished great success in a diversity of applications across many disciplines. In current work, we have combined two novel neural networks for classifying liver lesions of Hepatocellular carcinoma, Metastases, Hemangiomas, and Healthy tissues. On the first model called FireNet, we have introduced fire modules to reduce the model size and the number of parameters for quick classification. The first part will be in charge of extracting spatial feature information. The second model called Modified Long Short-Term Memory (MLstm), the features extracted by the FireNet are then used for temporal information and prediction with a new loss function called G-loss. In order to improve the proposed FireNet-MLstm model, a new bias was added through the forget gate. The activation function and the hyperbolic tangent were also added, which increased prediction accuracy. The dataset used in this study was collected by searching for the medical records with HCC, MET, HEM, and Healthy tissues in Jiangbin Hospital, the affiliated hospital of Jiangsu University from 2015 to 2018, with 120 patients, 30 patients with one or multiple Hcc, 26 patients with one or multiple Hem, 23 patients with one or multiple Met and 41 Healthy with non-lesion. The final classification accuracy of the proposed FireNet-MLstm model was 91.2%.

Keywords Deep learning · Classification of liver lesions · Computed tomography (CT) · Convolutional neural networks · Long short-term memory networks (LSTMs)

1 Introduction

According to the statistics of the World Health Organization, the second main cause of death in the world is liver cancer, out of all the cancers, and it is one of the leading

✉ Gedeon Kashala Kabe
gedeonkashala26@gmail.com

Yuqing Song
yqsong@ujs.edu.com

Zhe Liu
lzhe@ujs.edu.cn

[1] School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, China

causes of death worldwide [24, 27]. Medical image analysis plays a vital role in early diagnosis, and is an active way to decrease cancer mortality [2, 34]. In medical imaging task, the radiologist or other physicians can easily determine the lesion areas, but it is difficult to classify the lesions as benign and malignant. Presently, although computed tomography (CT) is often used to assist liver tumors diagnosis, the amount of information collected from computed tomography (CT) is enormous, and it is difficult for radiologists to interpret all the images in a short time [7, 18, 19]. In order to overcome this challenge, radiologists' benefit from computer-assisted decision systems. In recent years, deep learning methods, and in particular convolutional neural networks (CNN) have demonstrated an outstanding performance in image classification and visual object recognition in imagery [1, 6, 32]. Basically, small datasets are more prone to overfitting than large datasets. To avoid overfitting while training, researchers have used many methods, including data augmentation [22], Dropout [31], Rectified Linear Unit (ReLU) for activation [26] or transfer learning. Some researchers have tried to modify networks by decreasing the number of parameters with state-of-the-art performance while preserving accuracy [17]. SqueezeNet [15] is a smaller CNN architecture that executes better performance than AlexNet with 50× fewer parameters. Recurrent neural network (RNN), is based on the concept of allowing previous outputs to be used as inputs to the current step. Long short-Term Memory networks (LSTMs) are a special kind of RNN, capable of learning long-term dependencies [8, 30]. More researchers have been applying CNN–RNN methods to solve the problem of image classification. Guo et al. [11] propose to utilize the CNN–RNN framework to address the hierarchical image classification task. CNN allows them to obtain discriminative features for the input images, and RNN enables them to jointly optimize the classification of coarse and fine labels. In [28] Boon et al. they introduced a cascade of CNN with Long Short-Term Memory (LSTM) Network for classification of 3D brain tumor MR images into HG and LG glioma. Features from pre-trained VGG-16 were extracted and fed into LSTM network for learning high-level feature representations to classify the 3D brain tumor volumes into HG and LG glioma. In [20] Liu presented a novel approach of feature selection and dimension reduction for high-dimensional DNA microarray data.

This study extends our previous work on optimization of FireNet for liver lesion classification [17]. We present a novel approach for classifying liver lesions of Hepatocellular carcinoma (HCC), Metastases (MET), Hemangiomas (HEM), and Healthy tissues. The main idea of our research is to construct a succinct model with fewer parameters while preserving accuracy. Our model is a 2 stages approach, (1) We have used convolutional neural networks (CNN) with a succinct model called FireNet to realize the classification of 4 types of liver lesions. FireNet reduces the model size and the number of parameters by using fire modules from SqueezeNet while improving speed for quick classification. FireNet will be in charge of extracting spatial feature information. (2) The output of FireNet will be used in MLstm as input for sequence prediction. Our contributions to this work can be summarized as follows:

- We have combined two novel networks named FireNet-MLstm for classifying Liver Lesions.
- We have used fire modules to decrease the model size and the number of parameters for a concise model, which could save time in computer aided diagnosis of liver lesions.
- We have used MLstm to handle a sequence of features by adding a bias of 1 through the forget gate. In this part, the activation function and the hyperbolic tangent were added to improve the proposed FireNet-MLstm model.

– We have proposed a new loss function named G-loss to evaluate the parameters of the model.

The rest of the study is organized, as follows: In Sect. 2, we present an overview of the related works. In Sect. 3, we introduce our proposed method. We present our experiment and results in Sect. 4. We present the conclusions of our work in Sect. 5.

2 Related work

Recently there are several proposed approaches for classifying liver lesion or liver tumor by using deep neural network methods.

In 2014, Wu et al. have proposed the model which uses the deep learning method for the differentiation between the benign and malignant focal liver lesions. As a result, it has been found that the proposed method performed better than the available methods. The only problem is that method has a very small dataset which was not enough for the determination of the performance of the model [33].

In [10] the authors presented methods for generating synthetic medical images using recently presented deep learning Generative Adversarial Networks (GANs). The classification performance using only classic data augmentation yielded 78.6% sensitivity and 88.4% specificity. By adding the synthetic data augmentation, the results increased to 85.7% sensitivity and 92.4% specificity.

In 2016, Ben-Cohen et al. have applied the fully convolutional neural network for the segmentation purpose of the liver tumor. They have applied it for a small dataset and compared it with a patch-based convolutional neural network. Though the accuracy was 80%, it was tested only on a small dataset [3].

In 2017, Han has developed a deep Convolutional Neural Networks. This model used to work in 2.5D by taking the stack for the adjacent slices as the input and then producing the segmentation result according to the center slice. Being the best model of that time, there was still one disadvantage, i.e., there is no clarification about the proper optimal network architecture and also it was time-consuming and less accurate [13].

In [9] I. Diamant et al. described a novel method for automated diagnosis of liver lesions in portal phase computed tomography (CT) images that improves over single-dictionary BoVW methods by using an image patch representation of the interior and boundary regions of the lesions. The classification accuracy was 99%, respectively, and 93% for a combined dataset.

In 2017, Hoogi et al. have proposed an adaptive model for the convolutional neural network. They have verified their model for a dataset of reasonable size, i.e., 164 MRI and 112 CT images of liver lesions. For all the cases, they have evaluated their model for the various parameters like Dice similarity coefficients. It has been found that this model was significantly accurate than the other present models [14].

Özyurt et al. [25] presented a model capable to classify raw computed tomography (CT) images with CNN and obtained a 94.6% accuracy rate.

Kabe et al. [17] have proposed a classification of four types of liver lesions, namely, hepatocellular carcinoma, metastases, hemangiomas, and healthy tissues using convolutional neural networks with a succinct model called FireNet. They improved speed for quick classification and decreased the model size and the number of parameters by using fire modules from SqueezeNet. They have proposed a new Particle Swarm

Optimization (NPSO) to optimize the network parameters in order to further boost the performance of the proposed FireNet. The experimental results show that the parameters of FireNet are reduced 9.5 times smaller than GoogLeNet, 51.6 times smaller than AlexNet, and 75.8 smaller than ResNet. The size of FireNet is reduced 16.6 times smaller than GoogLeNet, 75 times smaller than AlexNet and 76.6 times smaller than ResNet. The final accuracy was 89.2%.

3 Proposed model

Our proposed model consists of mainly two parts, first part will be in charge of extracting feature with the aim to reduce the model size and the number of parameters for quick classification. Second part is for temporal feature extraction and prediction with a new loss function named G-loss. In order to improve the classification accuracy of FireNet-MLstm, the activation function and the hyperbolic tangent were added to the second part of proposed model to optimize the final results. The overall structure of the proposed FireNet-MLstm method for liver lesions classification is shown in Fig. 1.

3.1 Spatial feature extraction using FireNet

In this part of spatial feature extraction, FireNet introduces 8 fire modules to reduce the size model and the number of parameters during the training. Here the aim is to replace traditional layers and getting a model with a fewer parameters. We divide the layers of FireNet into three blocks. The first block contains fire2 and fire3; fire4 and fire5 belong to second block; fire6 to fire9 belong to third block. In block1, the output of fire2 is connected to the output of fire3 as the input of fire4. In block2, concatenation is applied after fire5. In block3, concatenation is applied after fire6 and fire8. Fire module contains two layers: a squeeze layer and an expand layer. Squeeze layer helps to decrease

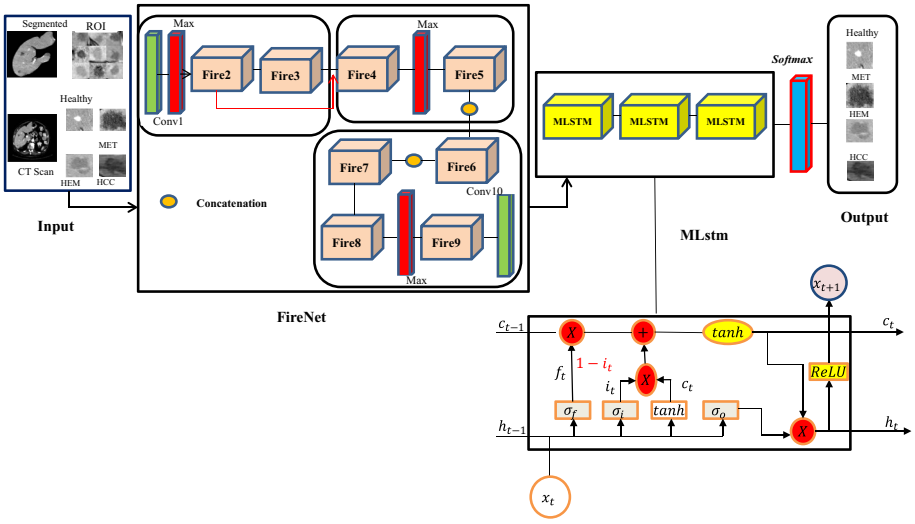


Fig. 1 Overall structure of our proposed FireNet- MLstm method for liver lesions classification

the number of input channels to 3×3 filters [15]. To maintain a small total number of parameters in our model, we need also to decrease the number of 3×3 filters. In squeeze layer, we used 1×1 filters and in expand layer, we used a combination of 3×3 filters and 1×1 filters. The input passes through the squeeze layer, then the data passes through the expand layer and depth of the data is expanded to conv 3×3 of the output tensor depth. To avoid any confusion between input channel number and output channel number, we have added convolution layer of 1×1 above of each concatenation. In this paper, we have started with a single convolutional layer as conv1 96, we have added the batch normalization before the nonlinearity to reduce internal covariate shift and overfitting, followed by eight fire modules including fire2 (96,16,64), fire3 (128,16,64), fire4 (128,32,128), fire5 (256,32,128), fire6 (256,48,192), fire7 (384,48,192), fire8 (384,64,256), fire9 (512,64,256), ending with a last convolutional layer conv10 512. The max-pooling with a stride of two is performed before fire2, fire5, and fire9. In this step of spatial feature extraction, we have used three kinds of kernels. First, in convolutional layer (conv1), the kernel size is 3×3. We have used it as a kernel with a large size. Second, the kernel size is 2×2; we have used them for extracting high-dimensional semantic information. Third, kernel size is 1×1; we have used them in squeeze layers and expand layers for extracting more useful information and discarding redundant information. FireNet is employed as the spatial features extraction with input of size of 64×64x1 which speed up the process of features extractions and enable us to get an output of size of 512,4 Feature map. We have applied the method of concatenation for well preserving the feature information of different layers used for classification. Our FireNet connections can be calculated as follows:

$$X_n = S\left(x_{n-1}\right) + x_{n-1} \tag{1}$$

where $x_{n-1}$where denotes the input of the first fire module, S denotes a nonlinear function. When the method of concatenation is applied, our FireNet connections can be calculated as follows:

$$X_n = S\left(\left[m\left(x_{n-1}\right), m\left(x_{n-2}\right), m\left(x_{n-3}\right)\right]\right) \tag{2}$$

where m is a 1×1 convolution. The structure of the FireNet is illustrated in Fig. 2.

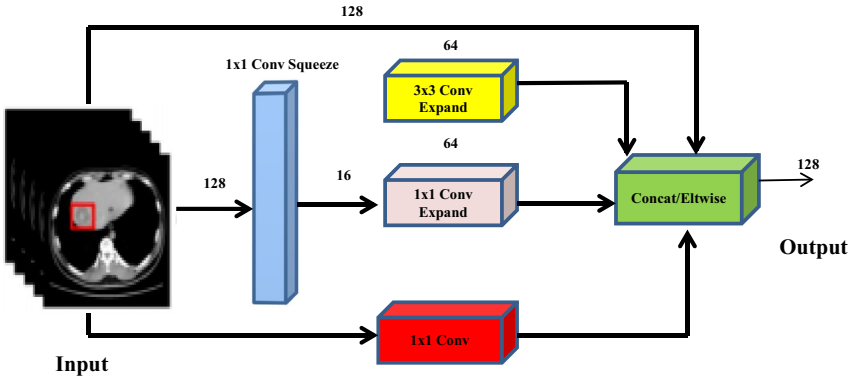


Fig. 2 Flowchart of FireNet for concatenation

3.2 Temporal feature extraction using MLstm

At this stage, the output of the FireNet is passing through MLstm. We have used MLstm to capture the temporal information from the features extracted used in the FireNet [29]. In standard Lstm architecture, the main reason for missing out some use full information is to delete a quantity of past information and to add new information separately [35], but in this study, we have modified Lstm by calculating the amount of information based on the amount of new information required to add. $\boldsymbol{f_t} = (1 - \boldsymbol{i})$, is the amount of new information required to add in our case. In addition, we have deleted negative outputs by adding the activation function after the output gate and the hyperbolic tangent was added which increased prediction accuracy. The structure of MLstm is shown in Fig. 3.

The equations for the gates in our MLstm are:

$$I_t = \sigma\left(X_t * W_{x_i} + H_{t-1} * W_{h_i} + b_i\right) \tag{3}$$

$$f_t = 1 - \mathrm{i_t}$$

$$g_t = \tanh\left(X_t * W_{x_g} + H_{t-1} * W_{hg} + b_g\right)$$

$$O_t = \sigma\left(X_t * W_{x_o} + H_{t-1} * W_{x_o} + b_o\right)$$

$$C_t = X_{t-1} * f_t + g_t * I_t$$

$$h_t = \tanh\left(c_t\right) * O_t$$

In Eq. (3), $\mathrm{I_t}$ denotes the input gate; σ is the sigmoid function; $\mathrm{X_t}$ denotes the input values; $\mathrm{W_{xi}}$ denotes the weight values of the artificial neural network; $\mathrm{H_{t-1}}$ denotes the output values received from the previous cell; $\mathrm{W_{hi}}$ is the weight matrix for output gate; $\mathrm{b_i}$ is the bias weight values. $\mathrm{f_t}$ is the forget gate; $1 - \mathrm{I_t}$ is a bias added through the forget gate, which modified the input gate. $\mathrm{g_t}$ denotes the variables; $\mathrm{O_t}$ denotes the output gate; $\mathrm{C_t}$ denotes the memory cell; $\mathrm{h_t}$ denotes the previous cell hidden state. If the output of the sigmoid function is 0, then the information is totally forgotten. If the output is 1, then the information is entirely

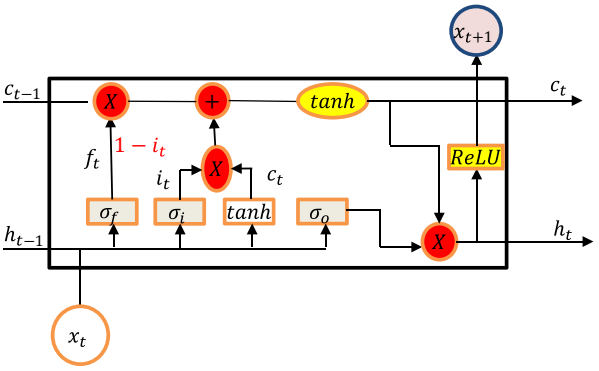


Fig. 3 MLstm (with added a bias of 1-I to forget gate and added ReLU, which increased prediction accuracy)

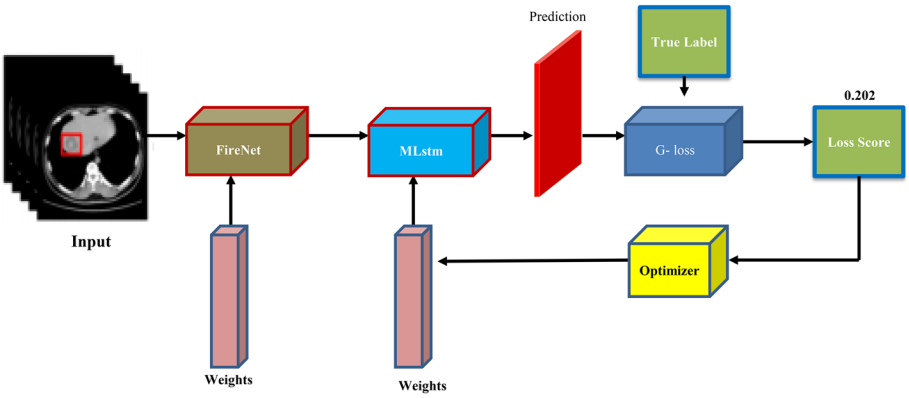


Fig. 4 G-loss Architecture

remembered. Our MLstm network comprises of sequence input with one dimension, 100 hidden units, fully connected layer, Softmax classifier, and classification output layers.

3.3 Loss function G-loss

Loss function is a method by which an optimizer is attempting to reduce the error in the prediction [16]. In this work, we have proposed a new loss function named G-loss for multi-class. Inspired by Cross-entropy function [12] and focal loss [21], we have added α as a hyper-parameter and $(1-y_i)^{\gamma}$ as a modulating factor to the C.E function. For multi-class, cross-entropy can be calculated as:

$$L_{CE} = -\sum_{i=1}^{M} l_i \log(y_i) \tag{4}$$

where M describes the number of classes, l_i describes a real probability distribution, y_i describes a probability distribution of the prediction. Where $l_i = 1$ if $i \in$ to the true label, else it is 0. We define the G-loss, as formulated in Eq. (3) bellow:

$$G_{Loss} = -\alpha.\frac{1}{M}\sum_{i=1}^{M}(1-y_i)^{\gamma} + (1-\alpha)l_i log(y_i) \tag{5}$$

where $(1-y_i)^{\gamma}$ denotes the modulating factor, γ denotes the focusing parameter, $\alpha \epsilon [0,1]$ is a hyper-parameter which control the behavior of our training. In this work, γ is used to regulate the rate at which easy examples are down-weighted. The focusing parameter γ is the main element by which the modulating factor is increased. When $\gamma = 0, G_{Loss}$ is equivalent to L_{CE}. In our experiments $\gamma = 2$, which works better. The structure of G-loss is illustrated in Fig. 4.

4 Experiments and results

4.1 Dataset

In this work, we have used the dataset of liver lesions collected from Jiangbin Hospital, the affiliated hospital of Jiangsu University (from 2015 to 2018) by searching for the medical

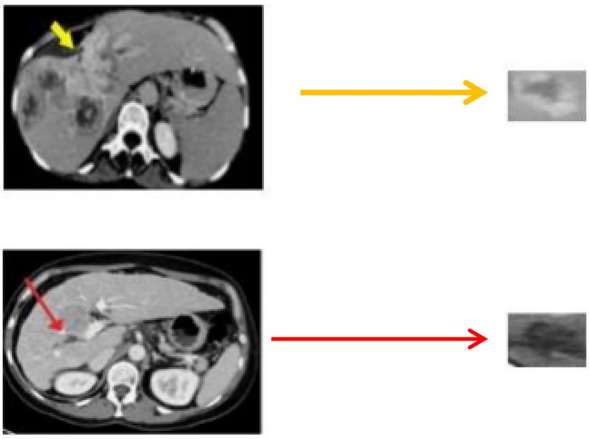


Fig. 5 ROI extraction process from a 2-D Computed tomography (CT) images of liver

records with HCC, HEM, MET, and healthy tissues. The dataset shows in Fig. 6 consists of a total of about 120 patients, 30 patients with one or multiple HCC, 26 patients with one or multiple HEM, 23 patients with one or multiple MET and 41 Healthy. The dataset had a total of 4142 images, including 1040 images of HCC, 1036 images of HEM, 1032 images of MET, and 1034 images of Healthy. From each class, we randomly choose 250 images for testing dataset, and the rest for training dataset. The margins in CT image dataset of liver lesions are shown in Fig. 5 that were used in this work were marked by an expert radiologist. Liver lesions are different in size, shape and contrast. We preprocessed in raw Digital Imaging and Communications in Medicine (DICOM) CT images 512 × 512-pixel metrics with a slice collimation of 5–7 mm, and 0.57–0.89 as an in-plane resolution range of slice. After all CT scan Hounsfield units (HU) values were truncated, and we normalized all slice intensities into the range [0,1] with min–max normalization. The result of each step of the proposed approach is shown in Fig. 7.

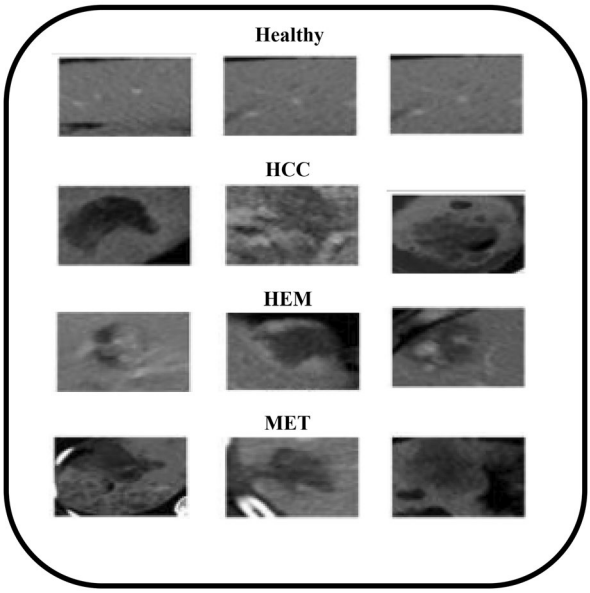
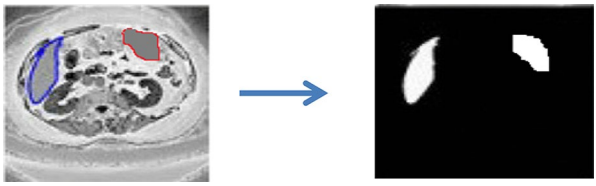


Fig. 6 Dataset examples

Fig. 7 Results from Dataset



4.2 Experiments

The proposed FireNet-MLstm model is trained on 75% of the training data, and tested on 25%. The model was coded in python using the pytorch framework. The experiments were performed using an NVIDIA GeForce GTX 980 Ti GPU. We have trained FireNet-MLstm model by using a learning rate of 0.0001 and stochastic gradient descent (SGD) with momentum of 0.9, Weight decay of 0.0001 for 175 epochs. Batch size is 64; the size of the input image is region of interests (ROIs) of 64×64 cropped from CT scans. We have first trained FireNet as a part of spatial feature extraction then we have used the features extracted by the FireNet for temporal information and prediction in our MLstm.

4.3 Results

We have used three types of measures to evaluate the performance of the presented model. These measures are precision, recall, accuracy, and F1 Score as expressed with true positive (TP), false positive (FP), true negative (TN), and false negative (FN) [36]. All these measures are described by the following equations:

$$TotalAccuarcy = \frac{(TP + TN)}{(TP + TN) + (FP + FN)} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

where TP denotes the number of image segments that were classified as having a type of lesion and had that type of lesion; TN denotes the number of image segments that were classified as having a type of non-lesion and had that type of non-lesion; FP denotes the number of image segments that were classified as having a type of lesion and had that type

Table 1 The results of the proposed method before to be combined

Method	Loss	Accuracy	Recall	Precision	F1-Score	Training time
SqueezeNet	0.382	81.8	74.7	75.9	75.2	2.4 s
Lstm	0.464	80.6	73.2	74.7	73.9	1.6 s
FireNet	0.327	82.7	80.1	78.1	79.0	2.0 s
MLstm	0.325	81.9	79.3	78.7	78.0	1.2 s

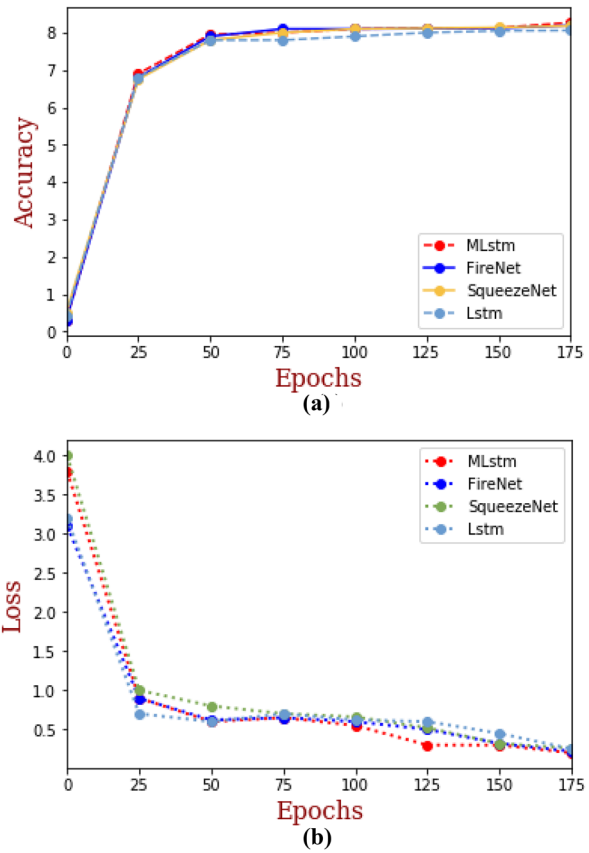


Fig. 8 Training progress of FireNet-MLstm before combination

of non-lesion; FN denotes the number of image segments that were classified as having a type of non-lesion and had that type of lesion.

Table 1 shows that the proposed model without combination had lower accuracy. In this stage, we train FireNet and MLstm without to be combined. One of the motivations to combine two networks is after getting the results are shown in Table 1.

Furthermore, Fig. 8 illustrates results of classification accuracies of four methods in liver lesions dataset. 8(a) shows the performance of the proposed FireNet-MLstm model without combination and other models using recall and precision. Also, the trend of these accuracies is shown in Fig. 8(b). From this figure the proposed FireNet-MLstm model

Table 2 The results of the proposed method with combination, number of parameters and model size

Method	Model size (MB)	Loss	Accuracy	Fire modules	Parameters	T/time (s)
SqueezeNet-Lstm	4	0.301	90.5	8	778,500	3.2
ResNet	230	0.298	90.1	–	60,344,115	5.1
GoogLeNet	50	0.276	90.4	–	7,424,344	4.8
FireNet-MLstm	3	0.202	91.2	8	730,890	2.3

T/Time = Training Time

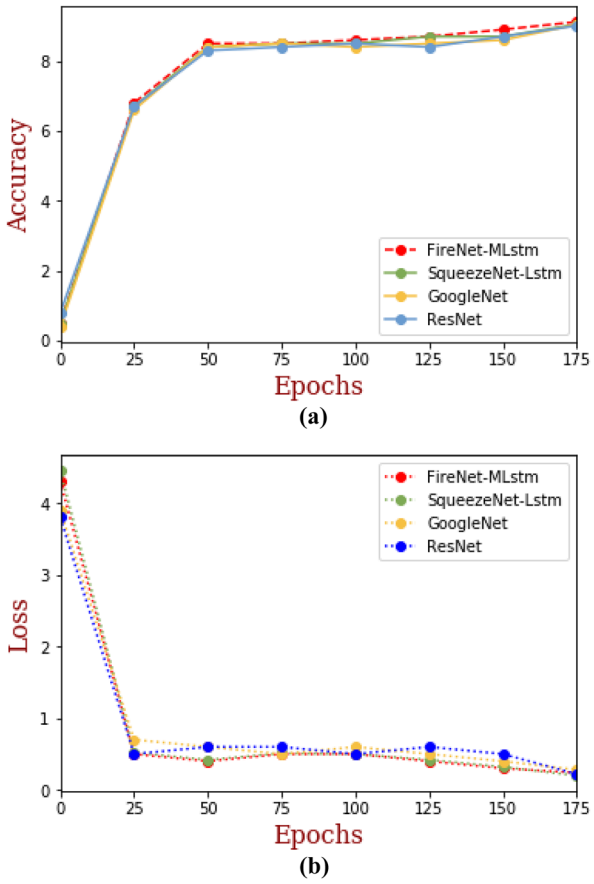


Fig. 9 Training progress of FireNet-MLstm with combination

without combination obtains the highest accuracy with 81.8% for FireNet and MLstm with 81.2% as compared to the SqueezeNet method and Lstm method. Also Fig. 8(c) shows the performance of the proposed FireNet-MLstm model without combination and other models during the training loss phase. The proposed FireNet-MLstm model network without combination, however, can drive the loss even lower than other models, and show a much faster rate of convergence.

Table 3 Comparison to state-of-the art methods

Method	Accuracy
Ben-Cohen et al. [3]	84
Hoogi et al. [14]	87.9
Frid-Adar et al. [10]	85.7
Kashala Kabe et al. [17]	89.2
Özyurt et al. [25]	94.6
Diamant et al. [9]	99.0
FireNet-MLstm	91.2

Table 4 The results of the proposed method with other classifiers

Method	Loss	Accuracy
SqueezeNet-MLstm	0.232	91.0
SqueezeNet-SVM	0.215	88.9
SqueezeNet-KNN	0.221	88.7
FireNet-Lstm	0.211	89.8
FireNet-SVM	0.208	90.6
FireNet-KNN	0.216	89.0
FireNet-MLstm	0.202	91.2

From Table 2, we have combined two novel neural networks called FireNet-MLstm by using 8 fire modules to decrease the number of parameters and model size. We applied concatenation method by adding a convolution 1×1 on top of each concatenation for well preserving the feature information of different layers used for classification. In addition, from MLstm, we have calculated the amount of information based on the amount of new information required to add. $\boldsymbol{f_t} = (1 - \boldsymbol{i})$, is the amount of new information required to add in our case. In addition, we have removed negative output by adding the activation function after the output gate and the hyperbolic tangent was added which increased prediction accuracy. The results show that the model size of FireNet-MLstm is decreased 1.3 times smaller than SqueezeNet-Lstm, 76.6 times smaller than ResNet, and 16.6 times smaller than GoogLeNet. The number of parameters of FireNet-MLstm is decreased to 1.06 times smaller than SqueezeNet-Lstm, 82.5 times smaller than ResNet, 10.1 times smaller than GoogLeNet. Table 2 shows that our model FireNet-MLstm performs better than standard CNN models and requires 2.3 s.

Figure 9a shows the performance of FireNet-MLstm after combination with the standards CNN models during the training phase. The results show that our proposed FireNet-MLstm model after combination works better than the standard CNN models. Figure 9b shows the performance of FireNet-MLstm after combination with the standards CNN models during the training loss phase. The results show that our proposed FireNet-MLstm model after combination works better than the other models.

Table 3 presents a comparison between the proposed method and the state-of-the art methods based on addressing the challenges in medical image classification, especially

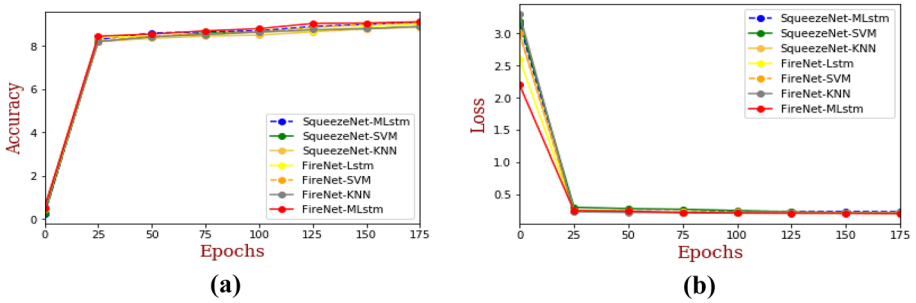


Fig. 10 Results of classification accuracies of proposed FireNet-MLstm with other methods in liver lesions dataset

Table 5 Performance of the proposed model with different training objectives evaluated using a 5-fold cross-validation scheme

Loss	Accuracy	Recall	Precision	FI-Score	Training Time
CE	89.2	87.9	88.2	88.0	2.7s
Focal loss	90.6	90.2	90.7	90.4	2.6s
MSE	89.8	89.1	88.1	88.5	2.8s
G-loss	91.2	91.1	90.7	90.8	2.3s

CE Cross entropy; *MSE* mean square error

liver lesions classification. The main challenge in medical imaging field is how to deal with the small datasets. Kabe et al. used the same dataset as this research. Our results presented a better performance than other methods.

Table 4 shows the results of the proposed FireNet-MLstm with other classifiers; MLstm, which is a widely used for improving the performance of classification applications, was applied by adding the amount of new information at forget gate; in addition, we have removed negative output by adding the activation function after the output gate and the hyperbolic tangent was added which increased prediction accuracy.

Furthermore, Fig. 10 illustrates results of classification accuracies of proposed FireNet-MLstm with other methods in liver lesions dataset. Also, the trend of these accuracies is shown in Fig. 10(a). From this figure the proposed FireNet-MLstm model obtains the highest accuracy with 91.2% as compared to other methods. Also Fig. 10(b) shows the performance of the proposed FireNet-MLstm model with other models during the training loss phase. The proposed FireNet-MLstm model network had better performance than the SVM and KNN classifiers.

We also evaluated the effect of the loss function on the model's prediction. To this end, we trained our model with stochastic gradient descent and used other widely used losses, the mean squared error (MSE), and the cross entropy (CE) which represents the dissimilarity of the approximated output distribution from the true distribution of labels. Training the model using CE yielded an average accuracy of 89.2, and 90.6 using focal loss, and 89.8 using MSE. As can be seen in Table 5, the proposed G-loss performed better than the other losses, most likely because this function is sensitive to both magnitude and angular variations in the output vectors.

Among the datasets, brain, pancreas, and public liver lesion yielded the best results; the performances of these datasets are reported in Table 6.

Table 6 Classification performance of the proposed model on different datasets

Dataset	Accuracy
Brain [23]	96.8
Pancreas [5]	96.7
Liver lesions [4]	97.3
Liver lesions (our dataset)	91.2

4.4 Discussion

In this study, we proposed FireNet-MLstm, a method for classifying liver lesions on CT images that can be explained as follow. The first stage is to apply FireNet to decrease the model size and the number of parameters by using fire modules from SqueezeNet. In this part, we used concatenation connections by adding a convolution layer of 1×1 on top of each concatenation for preserving the feature information of different layers used for classification. FireNet was in charge of extracting spatial feature from the image. Thus, the output of this part will be the input of our next step. In the second part, MLstm is used for temporal feature extraction. We modified the standard Lstm (MLstm) by adding a bias of 1 through the forget gate, the activation function (ReLU) was also added to remove negative outputs, which increased the accuracy. We have constructed a smaller model called FireNet-MLstm with few parameters based on the multi-class G-loss function with $\gamma = 2$, which can control the rate. Experimental results show that our model achieves better performances comparable with SqueezeNet-Lstm, GoogLeNet, and ResNet. The model size of FireNet-MLstm is decreased 1.3 times smaller than SqueezeNet-Lstm, 76.6 times smaller than ResNet, and 16.6 times smaller than GoogLeNet. The number of parameters of FireNet-MLstm is decreased to 1.06 times smaller than SqueezeNet-Lstm, 82.5 times smaller than ResNet, 10.1 times smaller than GoogLeNet. Table 2 shows that our model FireNet-MLstm performs better than standard CNN models and requires 2.3 s. These results prove the superiority of our proposed model with 91.2% accuracy.

5 Conclusions

In this paper, we proposed a new method that combines two novel neural networks for classifying liver lesions of Hepatocellular carcinoma, Metastases, Hemangiomas, and Healthy tissues. The first part of our method, named FireNet was in charge of spatial feature extraction by which we have introduced 8 fire modules to reduce the model size and the number of parameters for quick classification. We have added a 1×1 convolution layer above each concatenation to avoid any confusion between input channel number and output channel number for well preserving the feature information of different layers used for classification. The second part named MLstm was in charge of temporal information and prediction. We added a bias of 1 to the forget gate and the activation function (ReLU) was also added to remove negative outputs. This block of the model is coupled with G-loss function to evaluate the parameters. Experimental results show that the model size of FireNet-MLstm is decreased to 1.3 times smaller than SqueezeNet-MLstm, 76.6 times smaller than ResNet, and 16.6 times smaller than GoogLeNet. The number of parameters of FireNet-MLstm is decreased to 1.06 times smaller than SqueezeNet-Lstm, 82.5 times smaller than ResNet, 10.1 times smaller than GoogLeNet. Our model FireNet-MLstm presented a better performance than other and requires 2.3 s with 91.2% accuracy. The proposed model can help doctor's further diagnosis and avoid misdiagnosis. In the future work, we plan to improve the performance of our proposed method while increasing the number of CT images.

Acknowledgements The authors would like to thank Radiologists of the Medical Imaging department of Affiliated Hospital of Jiangsu University.

Author contributions GKK, YQS and ZL developed the main framework and collaborated in writing the paper. GKK Methodology and software; YQS and ZL supervision. All other authors contributed by revising the manuscript.

Funding This work was supported by the National Nature Science Foundation of China (61976106, 61772242, 61572239); China Postdoctoral Science Foundation (2017M611737); Six talent peaks project in Jiangsu Province (DZXX-122); Key special projects of health and family planning science and technology in Zhenjiang City (SHW2017019).

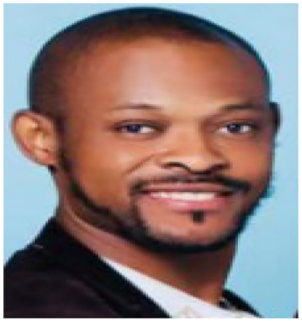
Declarations

Conflicts of interest The authors declare no conflict of interest.

References

1. Alahmer H, Ahmed A (2016) Computer-aided classification of liver lesions from CT images based on multiple ROI. Procedia Comput Sci 90:80–86. https://doi.org/10.1016/j.procs.2016.07.027
2. Bellver M, Maninis KK, Pont-Tuset J, Giró-i-Nieto X, Torres J, Van Gool L (2017) Detection-aided liver lesion segmentation using deep learning
3. Ben-Cohen A, Diamant I, Klang K, Amitai M, Greenspan H (2016) Fully convolutional network for liver segmentation and lesions detection.https://doi.org/10.1007/978-3-319-46976-8_9
4. Bilic P et al (2019) The Liver Tumor Segmentation Benchmark (LiTS). CoRR, vol. abs/1901.0. [Online]. Available: http://arxiv.org/abs/1901.04056
5. Chang Y et al (2017) Deep Learning based Nucleus Classification in Pancreas Histological Images
6. Chlebus G, Schenk A, Moltz JH, van Ginneken B, Hahn HK, Meine H (2018) Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing. Sci Rep 8(1):15497. https://doi.org/10.1038/s41598-018-33860-7
7. Christ PF et al (2016) Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, pp 415–423
8. Dai S, Li L, Li Z (2019) Modeling vehicle interactions via modified LSTM models for trajectory prediction. IEEE Access 7:38287–38296. https://doi.org/10.1109/ACCESS.2019.2907000
9. Diamant I et al (2016) Improved patch-based automated liver lesion classification by separate analysis of the interior and boundary regions. IEEE J Biomed Heal Informatics 20:1585–1594
10. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H (2018) GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing 321:321–331
11. Guo Y, Liu Y, Bakker EM, Guo Y, Lew MS (2018) CNN-RNN: a large-scale hierarchical image classification framework. Multimed Tools Appl 77(8):10251–10271. https://doi.org/10.1007/s11042-017-5443-x
12. Ho Y, Wookey S (2019) The real-world-weight cross-entropy loss function: modeling the costs of mislabeling. IEEE Access. https://doi.org/10.1109/ACCESS.2019.2962617
13. Han X (2017) Automatic liver lesion segmentation using a deep convolutional neural network method
14. Hoogi A, Subramaniam A, Veerapaneni R, Rubin DL (2017) Adaptive Estimation of active contour parameters using convolutional neural networks and texture analysis. IEEE Trans Med Imaging 36(3):781–791. https://doi.org/10.1109/TMI.2016.2628084
15. Iandola F, Han S, Moskewicz M, Ashraf K, Dally W, Keutzer K (2016) SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size
16. Janocha K, Czarnecki W (2017) On loss functions for deep neural networks in classification. Schedae Inform. https://doi.org/10.4467/20838476SI.16.004.6185
17. Kabe GK, Song Y, Liu Z (2020) Optimization of FireNet for liver lesion classification. Electronics 9(8):1237. https://doi.org/10.3390/electronics9081237
18. Li W, Jia F, Hu Q (2015) Automatic segmentation of liver tumor in CT images with deep convolutional neural networks. J Comput Commun 03:146–151. https://doi.org/10.4236/jcc.2015.311023
19. Li X, Chen H, Qi X, Dou Q, Fu C, Heng P (2018) H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Trans Med Imaging 37(12):2663–2674. https://doi.org/10.1109/TMI.2018.2845918

20. Li Z, Xie W, Liu T (2018) Efficient feature selection and classification for microarray data. PLoS ONE 13(8):1–21. https://doi.org/10.1371/journal.pone.0202167
21. Lin T, Goyal P, Girshick R, He K, Dollár P (2017) Focal Loss for Dense Object Detection. In: IEEE International Conference on Computer Vision (ICCV). pp 2999–3007. https://doi.org/10.1109/ICCV.2017.324
22. Miao S, Wang ZJ, Liao R (2016) A CNN regression approach for real-time 2D/3D registration. IEEE Trans Med Imaging 35(5):1352–1363. https://doi.org/10.1109/TMI.2016.2521800
23. Novak J et al (2021) Classification of paediatric brain tumours by diffusion weighted imaging and machine learning. Sci Rep 11(1):2987. https://doi.org/10.1038/s41598-021-82214-3
24. Organization WH (2018) World cancer report. World Health Organization, Geneva
25. Özyurt F, Tuncer T, Avci E, Koç M, Serhatlioğlu İ (2018) A novel liver image classification method using perceptual hash-based convolutional neural network. Arab J Sci Eng 44:1–10. https://doi.org/10.1007/s13369-018-3454-1
26. Pereira S, Pinto A, Alves V, Silva CA (2016) Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans Med Imaging 35(5):1240–1251. https://doi.org/10.1109/TMI.2016.2538465
27. Ryerson A et al (2016) Annual report to the nation on the status of cancer 1975–2012 featuring the increasing incidence of liver cancer. Cancer 122(9):1312–1337
28. Shahzadi I, Tang TB, Meriadeau F, Quyyum A (2018) CNN-LSTM: cascaded framework for brain tumour classification. In: IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), pp 633–637. https://doi.org/10.1109/IECBES.2018.8626704
29. Sherstinsky A (2020) Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D 404:132306
30. Smagulova K, James AP (2019) A survey on LSTM memristive neural network architectures and applications. Eur Phys J Spec Top 228(10):2313–2324. https://doi.org/10.1140/epjst/e2019-900046-x
31. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 15(1):1929–1958
32. Wang W et al (2018) Classification of focal liver lesions using deep learning with fine-tuning. In: Proceedings of the 2018 International Conference on Digital Medicine and Image Processing, pp. 56–60. https://doi.org/10.1145/3299852.3299860
33. Wu K, Chen X, Ding M (2014) Deep learning based classification of focal liver lesions with contrast-enhanced ultrasound. Optik (Stuttg) 125(15):4057–4063. https://doi.org/10.1016/j.ijleo.2014.01.114
34. Yasaka K, Akai H, Abe O, Kiryu S (2018) Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary Study. Radiology 286(3):887–896. https://doi.org/10.1148/radiol.2017170706
35. Wu Z, Wang X, Jiang YG, Ye H, Xue X (2015) Modeling spatial-temporal clues in a hybrid deep learning framework for video classification. pp 461–470.https://doi.org/10.1145/2733373.2806222
36. Zhu W, Zeng N, Wang N (2010) Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. NESUG Proc 19:67



Gedeon Kashala Kabe is a Ph.D. Scholar at school of Computer Science and Telecommunications Engineering of Jiangsu University. He received M.S degree in Computer Science and Technology from Harbin Engineering University in 2017, China. His current research pertains to the areas of Deep Learning, Medical Image Analysis and Computer Vision.

Yuqing Song received his Bachelor degree in medical sciences from Zhenjiang Medical College in 1981, his master degree in computer science and technology from Nanjing University of science and technology in 1999 and his Ph.D. degree in 2005 from Southeast University. He is now a full Professor at School of Computer Science and Telecommunication Engineering, Jiangsu University. Professor Song has published extensively in Chinese reviews and journals and won several provincial-level scientific awards. Currently, his research interests include medical image processing and analysis, datamining, and pattern recognition.

Zhe Liu got her Ph.D. degree in Computer Science in 2012 from Jiangsu University, Zhenjiang, China. She is a visiting scholar at the Department of Radiology at the University of Pittsburgh Medical Center, Pennsylvania, USA, and also an associate professor at the School of Computer Science and Telecommunication Engineering, Jiangsu University, Zhenjiang. Her research interests include image processing, data mining, and pattern recognition. She is a member of CCF and IEEE.