Deep learning neural network with transferlearning for liver cancer classification

Nibras Mizouri (nibras.elmizouri@gmail.com)
University of Sfax

Research Article

Keywords: Liver tumor, HCC, deep learning, transfer learning, VGG-16, MobileNet-V1

Posted Date: December 12th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2355564/v1

Additional Declarations: No competing interests reported.

Deep learning neural network with transfer learning for liver cancer classification

Nibras Mizouri[1,2*]

[1]MIRACL Laboratory, University of Sfax, Tunisia.
[2]Higher Institute of Computer Science and Multimedia, Sfax, Tunisia.

Corresponding author(s). E-mail(s): nibras.elmizouri@gmail.com;

Abstract

Hepatocellular carcinoma (HCC), "primary" cancer of the liver is the fourth most common cause of cancer related death worldwide. However,it can be reduced by early detection and diagnosis. With increasing use of Computed topography (CT) and Magnetic resonance (MR) imaging for diagnosis, traditional classification cancer manually is a difficult and time consuming task showing limitations in large and diverse datasets. Computer Aided Diagnosis (CAD) can play a key role in the early detection and diagnosis of liver cancer. Therefore the main objective of this work is to detect the liver cancer accurately using deep learning. We propose a novel CAD framework using convolution neural networks and transfer learning (pre-trained VGG-16 and MonileNet-V1 model). The classification accuracy of HCC reaches 96%. The system consists of three stages: (i) data collection , (ii) Data processing and (iii) Data Analysis.

Keywords: Liver tumor, HCC, deep learning, transfer learning, VGG-16, MobileNet-V1.

1 Introduction

Liver, the largest gland in the body, plays a vital and essential role to keep us alive. Its important functions are the storage of vitamins and nutrients, the production of bile essential for the digestions of fats,the filtration and the transformation of toxic substances in the blood. It also contributes to the storage of sugars and vitamins.

Liver cancer, ranked the 6th type of diagnosed cancer and the 3rd leading cause of death in the world in 2018 with around 841,000 new cases and 782,000 deaths per year [1]. Hepatocellular carcinoma (HCC)[2][3] is the most common type of primary liver cancer (accounting for 75% to 90% of cases). In most regions at high-risk of HCC (China, East Africa and Europe). The main determinants are infection with hepatitis B (HBV) or C (HCV), or cirrhosis of the liver caused by alcoholism. The incidence is higher in men than in women (the 5th cancer in men and the 9th in women) with a rate of 15.5% and 6.5% [4].

There are various modalities for diagnosing HCC liver lesions, such as CT scan and/or magnetic resonance imaging (MRI) with multiphase injection of contrast product, which are the two reference examinations that give precise images of the tumor and allow to see a possible extension to blood vessels of the liver [5].

However, a manual interpretation of a huge number of medical images can be a tedious and time consuming task, and can easily lead to errors for the radiologist when visually analyzing the image.

Consequently, since the 1980s, computer assisted diagnostic systems (CAD)[6][7] have been introduced to characterize liver damage. These systems which are based primarily on medical imaging and machine learning techniques, have proven high reliability and have become almost mandatory for use in clinical diagnosis, classification and treatment planning.

And as it is well known, Early detection of tumors plays a key role in the diagnosis of liver cancer and can improve long-term survival rates and avoid the risk of liver biopsy and surgery as well.

This helps automate the process so that a large number of cases can be treated, and provides quick and accurate results [8] especially with the digitization and immense progress in medical imaging as well as the massive improvement of artificial intelligence (AI) and machine learning in its "deep learning" version.

In this work, we wanted to present a robust CAD system for diagnosing and classifying liver tumours through a set of liver imaging based on deep learning and transfer learning. The rest of this document is organized as follow: Section 2 presents the related research. Section 3, presents the design of the proposed system that includes a detailed description of the HCC detection process. Section 4 treats the experimental results by analyzing the performance of the models used. Finally, Section 5 presents a overall conclusion of the study presented.

2 Related work

The term artificial intelligence brings together the sciences and technologies that can imitate, extend or even increase human intelligence using machines. In this regard, various automatic/semi-automatic techniques specialized in liver cancer diagnosis are used to improve procedures and remove diagnostic difficulties.

for example Chen et al.[9] proposed a CAD system for the classification of liver tumors based on GLCM texture functions. Huang et al.[10] presented a CAD based diagnostic approach to segmenting and classifying liver tumors using CT images. In their work, an accuracy of 81.7% was obtained using autocovariance texture characteristics [11].

Mala et al.[12] have developed a classification system for benign and malignant lesions using the biorthogonal wavelet transformation with the linear vector quantization (LVQ) neural network classifier. The system is tested with100 images composed of malignant (34 HCC, 18 cases of cholangio carcinoma) and benign (30 hemangioma and 18 cases of adenoma) lesions giving a perfermance rate of 92%.

Other methods optimized and effective, SVM based onliver cancer diagnosis [13]. Andrade et al.[14] have developed a classification model using the SVM (Support Vector Machines) classifier based on the following characteristics: firstorder gray level parameters (FOGLP), grey level co-occurrence matrix (GLCM), law texture energy (LTE), fractal dimension (FD) with an accuracy of 79.7%.

For[15], a liver tumor segmentation algorithm based on feature extraction and SVM classification to detect the tumor through the analysis of 150 CT images with 95% accuracy. In the same context, Wang et al.[16] apply an SVM classifier to establish a CAD system. This approach is based on first rate statistical texture characteristics and gray level co-occurrence matrix texture characteristics (GLCM) that can distinguish between normal livers and those attacked by HCC in CT images with an accuracy of 92.22%.

In recent years, deep learning has become the state of the art for various studies, since it allows significant advances in the treatment performance and classification of cancer tumors.In recent work, deep learning techniques using DNN have been successfully applied to solve a wide range of problems[17]. An accuracy of 99.38% was obtained using a deep neural net-work (DNN) [18] with a set of 225 images classified in three types hemangioma (HEM), hepatocellular carcinoma (HCC) and metastatic carcinoma (MET).

Kumar et al.[19] achieved 98% accuracy using the probabilistic neural network (PNN).

The Convoluted Neural Network (CNN)[20] has really caught the attention of many researchers attempting to solve this classification problem [21][22]. It is structured by a set of processing layers, especially the convoluted ones, as well as a set of activation functions by adding other elements such as the pooling layer and the fully connected layer.

Gibson et al[23] proposed a CNN based system to classify liver tumors from laparoscopic videos. Li et al.[24] presented a CNN algorithm applied to 30 images with an accuracy of 82.67%.

Another technique used in the field of medical imaging is transfer learning, [25] which involves fine tuning and reusing pretrained CNN models from a natural image datasets to new tasks. Reddy et al.[26] have chosen a CAD

system based on transfer learning with the VGG-16 model that reaches 90.6% an accuracy for the classification of normal liver images and steatosis.

Despite the success of these profound techniques, there are still inherent difficulties in their applicability.

First, there is the large amount of data needed to train the model. In addition, in medical imaging,this can be problematic because of the cost of acquisition,data anonymisation techniques, etc..

Second,massive data requires vast computing powers, and even using computing graphics units (GPUs). The training process takes a long time as well. Deep learning,is very resource-intensive. Therefore,each new proposal must take into account not only the performance, but also the calculation load.

In this article, we will be working on a problem of liver cancer (HCC) classification and detection while refining the restriction of having a small number of liver images per category in our database.

3 Materials and proposed methods

Figure 1. illustrates the operation of our CAD process for the diagnosis and detection HCC with transfer learning aid. The input data (liver MRI image) which undergoes a pre-processing stage, is subsequently used to form deep learning models.

The models performance is assessed using accuracy,sensitivity and specificity as presented in Section 4.

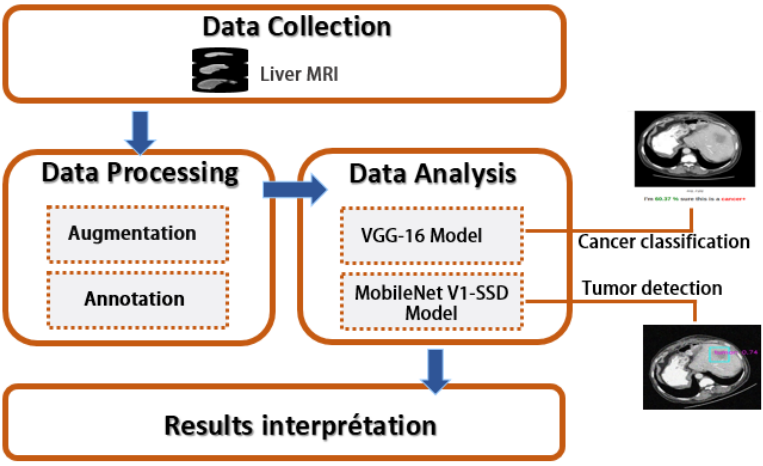

Fig. 1 Schema of proposed approach.

3.1 Data collection

The initial dataset consisted of 300 MRI images of liver acquired by physicians from patients. These MRI images were collected after injection of contrast material due to its superior contrast resolution than computer tomography (CT) and ultrasound; and its ability to provide both morphological and physiological information [27]. This base also presents benign (normal) and malignant (cancer) images of size (520 × 520) and (320 × 260) divided into two parts; 80% for training and 20% for the test.

3.2 Data Processing

Data Augmentation. It is generally known that training a large dataset gives better models of deep learning. However, collecting huge manually labeled data seems like a tough task The use of CNN architectures for the diagnosis and classification of liverless is based on the analysis of captured medical imaging of CT or MRI scans, both of which are expensive.

As a result, there is a challenge faced due to the inaccessibility of medical images due to patient confidentiality, manual and temporal demand and effort required by the radiologist. These problems have led to the technique of "data augmentation" [28] to solve the 'overfitting' problem, which presents an over learning model because of the limited base.The trained model suffers from being unable to predict the relevant areas on the new images. The increased data is used for improved classification performance [29][30].

A technique that creates updated copies of data, which means artificially increasing the size of the database by applying simple transformations to it while keeping the same format, the same size and length. We introduce these new images during the training which fulfills the following functions (Table 1).

Table 1 Data Augmentation Parameters

Parameter	value
Flip	0.3
Rotation	25°
Noise	0.5
Blur	0.1

We apply a variety of augmentation techniques such as rotation, translation, noise injection, and flip to produce new versions of existing images.

- Flip: the image by reversing rows or columns of pixels vertically or horizontally.
- Rotation: Rotate the image to the right or left on a 25°axis.
- Noise: Inject a matrix of random values (noise) usually drawn from a Gaussian distribution.
- Blur: blur an image using the average of a pixel and its neighbors.

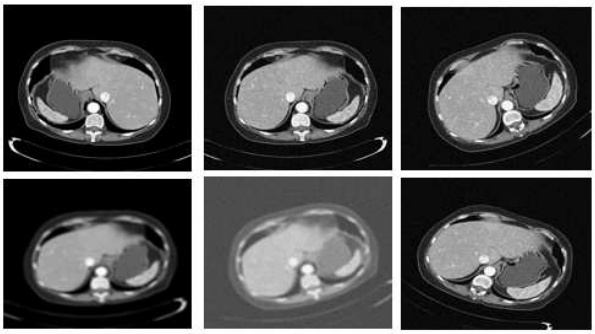

Fig. 2 Images augmentation with flip, rotation, noise and blur after applying DataSet Generator().

Data Annotation. Once the data quantity problem is resolved, an equally important element is needed in the process of our CAD system to classify and detect liver tumors. This element allows the data to be annotated by assigning a text or a caption that describes the image. In the medical field, assembling an annotated and labeled database presents a challenge because of limited resources.

In addition, it takes a lot of time during the acquisition which is rarely annotated.

Various studies have highlighted this technique [31] which aims to improve medical assistance by creating a help tool that describes, analyzes and locates such diseases in the con-tent of images. Good image annotations should be accurate in two directions. First, objects must be accurately labeled with the correct class. Second, pixels that contain the object must be accurately selected.

For this, there are three categories:the automatic,the manual and the semiautomatic [32][33]. Manual annotation consists of the user selecting the annotation form manually on paper. Whereas automatic annotation automatically detects and labels the semantic content of the images with a set of keywords. As for as semi-automatic annotation is concerned, it needs to integrate human assistance into the automatic annotation process.

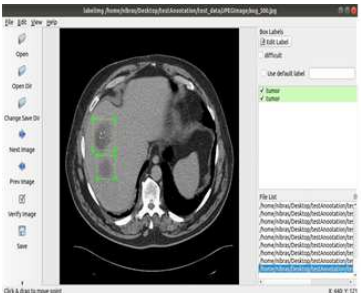

Fig. 3 Example image annotation using LabelImg..

In our approach, image annotation is used to detect and locate cancerous lesions. As an annotation tool, we choose "LabelImg" (Figure 3.) which has a graphical interface (in python) to annotate the image base by selecting the location of tumors. The resulting annotated base is saved as XML files which are divided into two subsystems training and testing for model formation.

3.3 Data Analysis using transfer learning

The second stage of our proposed CAD system is based on deep learning, a classifier based on transfer learning using two models VGG-16 and MobileNet-V1.

Transfer learning [34] is a powerful strategy that transfers knowledge acquired from solving one problem to deal with another differing model [35]. This method has been very successful with the rise of Deep Learning.

Indeed, the models used in this field require high computational times and significant resources. Using pretrained models as a starting point, Transfer Learning allows the rapid development of high-performance models and effectively solves complex problems.

Figure 4. Shows the contrast between the process of traditional machine learning and transfer learning. As we can see conventional machine learning, tries to learn each task separately with a different model from scratch. On the other hand transfer learning tries to extract knowledge from previous source tasks to target tasks where the latter has very little labeled data [36].

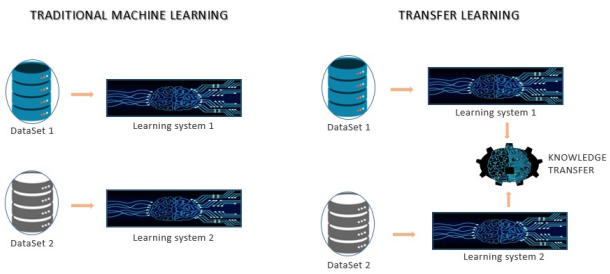

Fig. 4 Transfer learning vs. Machine learning.

Practically, the purpose of this study is to apply transfer learning with preformed CNN for the classification and detection of HCC tumors. This is based on a simple idea; to readapt the knowledge gained from the architecture of the CNN model (source) trained on a large mass of data to solve our target problem. We operate two different architectures VGG-16 and MobilNet.

VGG-16 Model . As shown in Figure 5., the architecture of the VGG-16 [37] model is pre-trained with transfer learning. It initially consists of several layers that take a 224×224 image from the ImageNet database (containing 14

miles of images)[38] and classifies it in one of 1000 classes, reaching an accuracy of 92.7%.

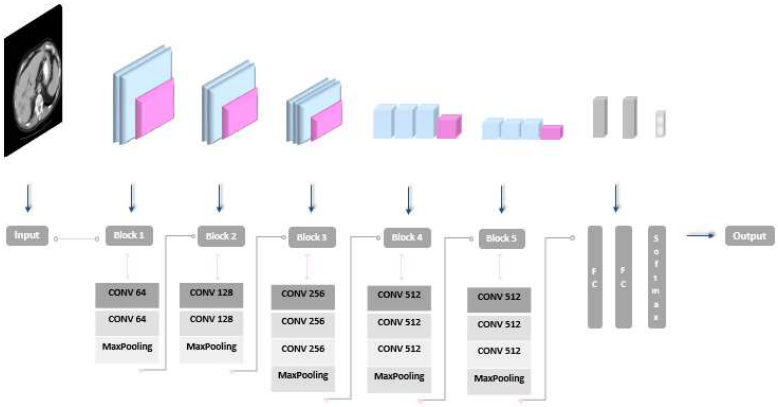

Fig. 5 VGG-16 architecture.

Our proposed VGG-16 classification model uses the model,which is simply by resetting the final layer fully connected by providing the required number of features and classes.

The first and second convolutional layers are followed by a max pooling layer, consisting of 64 filters. Taking as input a 224x224 input image (with depth3) passed through the first and second convolutional layer, the dimensions increase to 224x224x64.

Then, the third and fourth convolutional layers are 128 filters followed by a max pooling layer. The next three layers fifth,sixth and seventh are convolutional layers using 256 nodes with a pooling layer. Therefore, in the eighth to thirteenth convoluted layers the core size changes to 512 filters, followed by a maximum pooling layer.

Finally, the last block, consisting of two fully connected layers composed of 512 nodes and an output layer with two neurons,corresponds to the two classes (normal and cancer) for our study. The rectified linear unit (ReLU) is used as an activation function in all hidden layers, while the softmax activation function is used in the last layer of the CF layers.

MobileNet-V1-SDD Model . MobileNet is a special class of CNN proposed by Google Brain researchers in April 2017 [39]. The idea behind this model was to create a new type of convolution to the places of the standard convolutive layer; this new convolution is called “Depthwise separable convolution”.

This model introduces 2 global hyper-parameters: the width multiplier and t he resolution multiplier that can control the number of input/output channels of convolution layers and resolution of the input data (i.e. height, width).

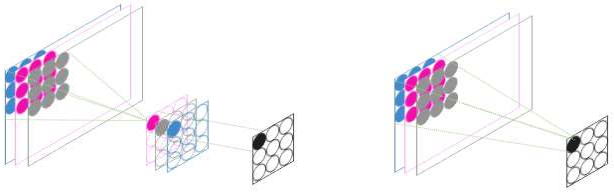

Fig. 6 The structural module: Depthwise separable convolution(Left) and Standard convolution(Right).

The "Depthwise separable convolution" consists of two separate layers: the first Depthwise Convolution and the Pointwise convolution [40][37] as listed in Figure 6. In this operation,deep convolution kernels apply a unique filter to each enter channel, followed by a Point convolution (1×1) to combine the product of the first convolution.

MobileNet is favored in the problem of detecting and locating objects in terms of computational cost and model size. Compared to standard convolution networks that filter and combine inputs into a new set of single step outputs, Deep Separable Convolution divides this operation into two layers, a separate layer for filtering and a layer for the combination.

Table 2 Calculation difference between depth separable and standard convolution

Convolution	Calculation cost
Standard	$D_k^2 * D_f^2 * M * N$
Depthwise Separable	$D_k^2 * D_f^2 * M + M * N * D_f^2$

Table 2. summarizes the cost formulas for the two methods, with D_k the size of the convolution kernel, D_f the size of the input layer and M, N represents, respectively, the number of channels of the input and output layer.

When we use the equations in Table 2., the reduction in computational costs and model parameters can be demonstrated by the following ratio:

$$\frac{\text{Cost}_{\text{separable}}}{\text{Cost}_{\text{conv}}} = \frac{1}{N} + \frac{1}{D_k^2} \tag{1}$$

Thus, through this relation, we find that the larger the size of the convolution kernel is the higher the number of output layers and the more interesting the convolution proposed in this paper is. If we set N to a large value and a classic filter size of 3, the cost of calculating by 1/9 is reduced by at least 88% [39].

Each convolution in the Mobilenet network is followed by a batch normalization [41]. Then the ReLU6 activation function [42] is applied and finally, a 1x1 convolution is used to obtain a map on which one still applies a batch normalization and a ReLU6 (figure7).

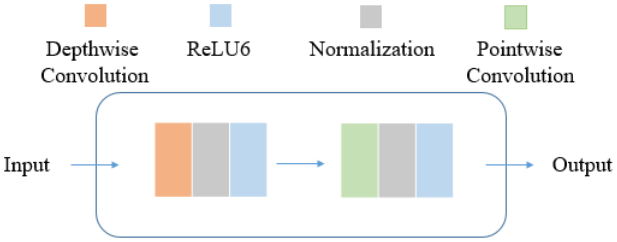

Fig. 7 Depthwise separable convolution block.

For our study, we use the MobileNet SSD combination for the detection and localization of liver tumors. The SSD (Single Shot Detector)[43] as a metadetector model is based on an auxiliary MobileNet network for our case, in order to obtain the location and category of the target in a regressive way.

It uses different feature maps, some of which come from the MobileNet core network, to perform location classification and regression. These feature maps have different sizes in order to take advantage of the high level and low-level information.

The operation of the MobileNet-V1 SSD is simple as explained above; the left side of the architecture (Figure 8) shows the MobileNet-V1 feature extraction backbone network with the addition of eight convolution layers behind the conv13. A total of 6 layers (4 layers and the conv11 and conv13 layers) are extracted for detection at the SSD level, and finally assigned to a detection unit.

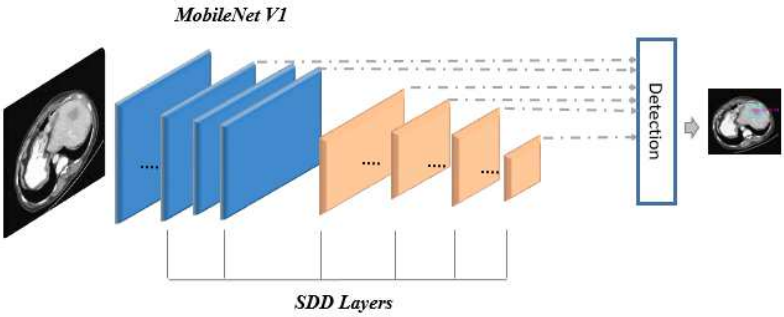

Fig. 8 MobileNet-V1-SSD network architecture.

4 Results and Discussion

We conducted all our tests and training on a laptop equipped with an Intel Core i7-7500U CPU / 3.50 Ghz processor with 8 GB RAM and an NVIDIA GeForce GT920M graphics card.

The CAD method presented, consists of two steps: the preparation of data and the entarinement of models used for the classification and detection of cancerous lesions. In the first step, two techniques are used: data augmentation and annotation which is used to build our datasets in order to improve the detection accuracy of the models that are trained.

In the second step, a VGG-16 model is applied to classify the images (cancer or normal), and another MobileNet-V1 SSD model used for the detection and localization of liver tumors found in cancerous livers.

With a total of 334 MRI images of the liver(Data Test), 267 images are used for training and 67 for testing. First, we will improve the accuracy by using the image enhancement technique that generates this set of images with the different transformations (section 3.2) resulting in a total of 2000 images. Finally, we will exploit the pre-trained model VGG-16 described in 3.3. to extract features and classify the images.

The analysis of the performance of the proposed model is defined by different criteria: Accuracy, Confusion Matrix, Precision, and F-score which are shown in Table 3.

The performance measures considered are defined as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{4}$$

$$F1 - \text{score} = 2 \times \frac{\text{Prcision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \tag{5}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

Where TP, FP, TN and FN indicate the number of true positives, false positives, true negatives and false negatives respectively.

Table 3 Evaluation metrics

Classifier	VGG-16
Accuracy(%)	96
Sensitivity(%)	94
Specificity(%)	98
Precision(%)	98
F1-score(%)	96

True positive refers to correctly identified tumor pixels, true negative refers to correctly identified non-tumor pixels, false positive refers to incorrectly identified tumor pixels and false negative refers to incorrectly identified non-tumor pixels.

All these measures are calculated for each class, and an overall measure of the algorithm is calculated by taking the average of all these measures across the two classes.

The proposed model is formed for 100 epochs. Figure 9. shows the accuracy of learning and testing at different times. Indeed, from Era 0 to Era 29, the accuracy curve for trains increases rapidly until it reaches 92%. Then, it converges to a value of 96%. A rapid decrease in the loss curve can be observed from epoch 0 to epoch 25 where the loss is equal to 0.12%; a kind of stability can then be observed up to the value of 0.09%.

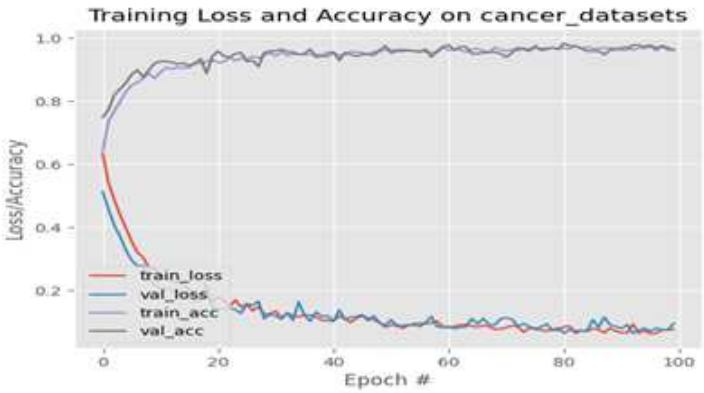

Fig. 9 Training and validation accuracy/loss of the proposed model VGG-16.

From the confusion matrix (Table 4), it can be observed that 173 of the 184 cases of hepatocellular carcinoma are correctly classified as cancer (TP), while 11 of them are erroneously classified as normal (FN). However, 179 of the 182 normal cases used for testing are correctly classified as normal (TN), but 3 of them are misclassified as hepatocellular carcinoma (FP).

Table 4 Confusion matrix for the classification model

		Predicted values	
		Cancer	Normal
Actual values	Cancer	173	11
	Normal	3	179

Before training the Mbilenet(v1) SSD detection model, an annotation step is necessary. The process is long and tedious but seems indispensable to obtain good results. It consists in drawing boxes around the tumors by giving them the label "tumor" with the LabelImg software.

The entire dataset (340 images annotated in .jpg and .xml) is used. About 80% of the imges of this dataset are used for the Train and 20% are used for the Test.

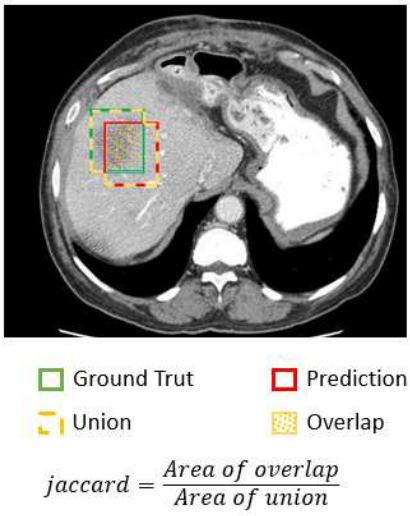

Fig. 10 Illustration of jaccard calculation.

The PASCAL VOC evaluation is used to evaluate the detector's localization and classification performance [44]. Based on the PASCAL VOC metric(Figure 10.), any detection with jaccard overlap 50% and an identical prediction label with groundtruth is considered a correct detection.

As the value FP, FN, and TP can be modified by setting different thresholds, it is important to evaluate the accuracy and sensitivity at different thresholds in order to measure the over-all performance of a detector model. Average Accuracy (AP) defines the average of the maximum accuracy at different recalls for each category and mAp is also defined as the average of the APs across all categories.

Practically, our used model MobileNet-V1-SSD is formed onthe basis of 430 images with 50 epochs, in order to learn howto predict the locations of liver tumors in a reduced time whichreaches 63% as mAp (Figure 11).

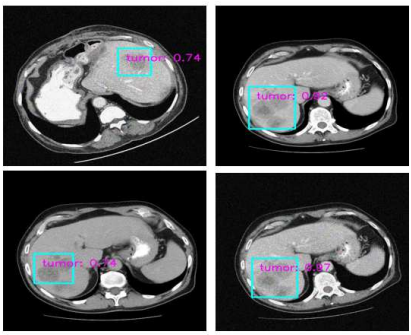

Fig. 11 Example cancer detection results using MobileNet SSD.

5 Conclusion

In this paper, we presented an automated CAD system for the detection of liver tumors (hepatocellular carcinoma) based on deep learning as well as transfer learning.

In the development phase, initially, we had a set of images augmentation and annotation techniques which were used to prepare the data. Then, a preformed model of VGG-16 as a classification model and a modified version of MobileNet-SSD was proposed as a tumor detection model.

The choice of models has been carefully selected to fit the nature of the material and limited resources. The analysis of the performance of the proposed system shows that the classification process produced an optimal accuracy of 96% with a negligible loss of validation.

This method is therefore, an efficient way to detect the cancerous area from liver images that will be useful for clinical diagnosis and decision making for early diagnosis.

Declarations

- **Conflict of interest** The corresponding author states that there is no conflict of interest.
- **Availability of data and materials** data not associated.

References

[1] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A.: Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians **68**(6), 394–424 (2018)

[2] Yang, J.D., Hainaut, P., Gores, G.J., Amadou, A., Plymoth, A., Roberts, L.R.: A global view of hepatocellular carcinoma: trends, risk, prevention

and management. Nature reviews Gastroenterology & hepatology, 1–16 (2019)

[3] Trinchet, J., Alperovitch, A., Bedossa, P., Degos, F., Hainaut, P., Van Beers, B.: Épidémiologie, prévention, dépistage et diagnostic du carcinome hépatocellulaire. Bulletin du cancer **96**(1), 35–43 (2009)

[4] Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D., Piñeros, M., Znaor, A., Bray, F.: Estimating the global cancer incidence and mortality in 2018: Globocan sources and methods. International journal of cancer **144**(8), 1941–1953 (2019)

[5] Burrel, M., Llovet, J.M., Ayuso, C., Iglesias, C., Sala, M., Miquel, R., Caralt, T., Ayuso, J.R., Solé, M., Sanchez, M., *et al.*: Mri angiography is superior to helical ct for detection of hcc prior to liver transplantation: an explant correlation. Hepatology **38**(4), 1034–1042 (2003)

[6] Van Ginneken, B., Romeny, B.T.H., Viergever, M.A.: Computer-aided diagnosis in chest radiography: a survey. IEEE Transactions on medical imaging **20**(12), 1228–1241 (2001)

[7] Goceri, E., Shah, Z.K., Layman, R., Jiang, X., Gurcan, M.N.: Quantification of liver fat: a comprehensive review. Computers in biology and medicine **71**, 174 (2016)

[8] Goryawala, M., Guillen, M.R., Cabrerizo, M., Barreto, A., Gulec, S., Barot, T.C., Suthar, R.R., Bhatt, R.N., Mcgoron, A., Adjouadi, M.: A 3-d liver segmentation method with parallel computing for selective internal radiation therapy. IEEE Transactions on Information Technology in Biomedicine **16**(1), 62–69 (2011)

[9] Chen, E.-L., Chung, P.-C., Chen, C.-L., Tsai, H.-M., Chang, C.-I.: An automatic diagnostic system for ct liver image classification. IEEE transactions on biomedical engineering **45**(6), 783–794 (1998)

[10] Huang, Y.-L., Chen, J.-H., Shen, W.-C.: Computer-aided diagnosis of liver tumors in non-enhanced ct images. Journal of Medical Physics **9**, 141–150 (2004)

[11] Huang, Y.-L., Chen, J.-H., Shen, W.-C.: Diagnosis of hepatic tumors with texture analysis in nonenhanced computed tomography images. Academic radiology **13**(6), 713–720 (2006)

[12] Mala, K., Sadasivam, V.: Wavelet based texture analysis of liver tumor from computed tomography images for characterization using linear vector quantization neural network. In: 2006 International Conference on Advanced Computing and Communications, pp. 267–270 (2006). IEEE

[13] Zhou, J., Xiong, W., Tian, Q., Qi, Y., Liu, J., Leow, W.K., Han, T., Venkatesh, S.K., Wang, S.-c.: Semi-automatic segmentation of 3d liver tumors from ct scans using voxel classification and propagational learning. In: MICCAI Workshop, vol. 41, p. 43 (2008)

[14] Andrade, A., Silva, J.S., Santos, J., Belo-Soares, P.: Classifier approaches for liver steatosis using ultrasound images. Procedia Technology **5**, 763–770 (2012)

[15] Rajagopal, R., Subbiah, P.: Computer aided detection of liver tumor using svm classifier. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering **3**(6), 10170–7 (2014)

[16] Yu, L., Wang, C., Cheng, S., Guo, L.: Establishment of computer-aided diagnosis system for liver tumor ct based on svm. In: 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), pp. 710–715 (2018). IEEE

[17] Sun, C., Guo, S., Zhang, H., Li, J., Chen, M., Ma, S., Jin, L., Liu, X., Li, X., Qian, X.: Automatic segmentation of liver tumors from multiphase contrast-enhanced ct images based on fcns. Artificial intelligence in medicine **83**, 58–66 (2017)

[18] Das, A., Acharya, U.R., Panda, S.S., Sabut, S.: Deep learning based liver cancer detection using watershed transform and gaussian mixture model techniques. Cognitive Systems Research **54**, 165–175 (2019)

[19] Kumar, S., Moni, R., Rajeesh, J.: An automatic computer-aided diagnosis system for liver tumours on computed tomography images. Computers & Electrical Engineering **39**(5), 1516–1526 (2013)

[20] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998)

[21] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Medical image analysis **42**, 60–88 (2017)

[22] Shin, H.-C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, R.M.: Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. IEEE transactions on medical imaging **35**(5), 1285–1298 (2016)

[23] Gibson, E., Robu, M.R., Thompson, S., Edwards, P.E., Schneider, C.,

Gurusamy, K., Davidson, B., Hawkes, D.J., Barratt, D.C., Clarkson, M.J.: Deep residual networks for automatic segmentation of laparoscopic videos of the liver. In: Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling, vol. 10135, p. 101351 (2017). International Society for Optics and Photonics

[24] Li, W., *et al.*: Automatic segmentation of liver tumor in ct images with deep convolutional neural networks. Journal of Computer and Communications **3**(11), 146 (2015)

[25] Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Transactions on knowledge and data engineering **22**(10), 1345–1359 (2009)

[26] Reddy, D.S., Bharath, R., Rajalakshmi, P.: A novel computer-aided diagnosis framework using deep learning for classification of fatty liver disease in ultrasound imaging. In: 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom), pp. 1–5 (2018). IEEE

[27] Mikołajczyk, A., Grochowski, M.: Data augmentation for improving deep learning in image classification problem. In: 2018 International Interdisciplinary PhD Workshop (IIPhDW), pp. 117–122 (2018). IEEE

[28] Taha, B., Dias, J., Werghi, N.: Classification of cervical-cancer using pap-smear images: a convolutional neural network approach. In: Annual Conference on Medical Image Understanding and Analysis, pp. 261–272 (2017). Springer

[29] Maayan, F., Eyal, K., Jacob, G., Hayit, G.: Gan-based data augmentation for improved liver lesion classification. arXiv preprint (2018)

[30] Abd Manaf, S., Nordin, M.J.: Review on statistical approaches for automatic image annotation. In: 2009 International Conference on Electrical Engineering and Informatics, vol. 1, pp. 56–61 (2009). IEEE

[31] Barnard, K., Shirahatti, N.V.: Method for comparing content-based image retrieval methods. In: Internet Imaging IV, vol. 5018, pp. 1–8 (2003). International Society for Optics and Photonics

[32] Wenyin, L., Dumais, S.T., Sun, Y., Zhang, H., Czerwinski, M., Field, B.A.: Semi-automatic image annotation. In: Interact, vol. 1, pp. 326–333 (2001)

[33] Torrey, L., Shavlik, J.: Transfer learning. In: Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, pp. 242–264. IGI global, ??? (2010)

[34] Sarkar, D., Bali, R., Ghosh, T.: Hands-On Transfer Learning with Python: Implement Advanced Deep Learning and Neural Network Models Using TensorFlow and Keras. Packt Publishing Ltd, ??? (2018)

[35] Weiss, K., Khoshgoftaar, T.M., Wang, D.: A survey of transfer learning. Journal of Big data **3**(1), 9 (2016)

[36] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)

[37] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Communications of the ACM **60**(6), 84–90 (2017)

[38] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 (2017)

[39] Sifre, L., Mallat, S.: Rigid-motion scattering for image classification. Ph. D. thesis (2014)

[40] Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)

[41] Krizhevsky, A., Hinton, G.: Convolutional deep belief networks on cifar-10. Unpublished manuscript **40**(7), 1–9 (2010)

[42] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C.: Ssd: Single shot multibox detector. In: European Conference on Computer Vision, pp. 21–37 (2016). Springer

[43] Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International journal of computer vision **88**(2), 303–338 (2010)

[44] Bruegel, M., Rummeny, E.J.: Hepatic metastases: use of diffusion-weighted echo-planar imaging. Abdominal imaging **35**(4), 454–461 (2010)