

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,600

Open access books available

178,000

International authors and editors

195M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



## Chapter

# Deep Learning Techniques for Liver Tumor Recognition in Ultrasound Images

*Delia Mitrea, Sergiu Nedevschi, Mihai Socaciu and Radu Badea*

## Abstract

Cancer is one of the most severe diseases nowadays. Thus, tumor detection in a non-invasive and accurate manner is a challenging subject. Among these tumors, liver cancer is one of the most dangerous, being very common. Hepatocellular Carcinoma (HCC) is the most frequent malignant liver tumor. The golden standard for diagnosing HCC is mainly the biopsy, however invasive and risky, leading to infections, respectively to the spreading of the tumor through the body. We conceive computerized techniques for abdominal tumor recognition within medical images. Formerly, traditional, texture-based methods were involved for this purpose. Both classical texture analysis methods, as well as advanced, original texture analysis techniques, based on superior order statistics, were involved. The superior order Gray Level Cooccurrence Matrix (GLCM), as well as the Textural Microstructure Cooccurrence Matrices (TMCM) were employed and assessed. Recently, deep learning techniques based on Convolutional Neural Networks (CNN), their fusions with the conventional techniques, as well as their combinations among themselves, were assessed in the approached field. We present the most relevant aspects of this study in the current paper.

**Keywords:** hepatocellular carcinoma (HCC), ultrasound images, deep learning techniques, convolutional neural networks (CNN), conventional machine learning (CML), classification performance assessment

## 1. Introduction

Cancer is one of the most severe and frequent affections nowadays, being lethal in most situations. In particular, liver cancer incidence has considerably increased during the last years, from 841,000 cases in 2018, to 905,700 cases in 2020, the number of cases being estimated to double by 2040. HCC is the most often met liver cancer, present in 70% of the primary hepatic cancer cases, being the 4th most frequent liver malignant tumor in men, the 7th most frequent liver malignant tumor in women, respectively the 3rd most frequent cancer-related cause of death after lung and colorectal cancer. HCC usually evolves from cirrhosis, after a liver parenchyma restructuring phase [1]. The best method for diagnosing liver cancer is through

biopsy, which raises risks, conducting to infections, respectively to the spreading of the tumor in the human body [2]. Thus, advanced, computerized methods are due, for revealing subtle aspects within medical images, respectively for achieving a both non-invasive and highly accurate diagnosis process. Ultrasonography is a medical image-based investigation technique that is non-invasive, unexpensive, and suitable for patient disease monitoring. Other alternative medical examination techniques, such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) could be irradiating and/or expensive. So, in our research, we performed computer aided and automatic diagnosis of HCC based on ultrasound images. Firstly, classical, as well as advanced texture analysis techniques were employed to perform a refined differentiation between the tumoral and the non-tumoral tissue. Superior order Generalized Cooccurrence Matrices (GCM), in the form of the Gray Level Cooccurrence Matrix (GLCM) of superior order, respectively Textual Microstructures Cooccurrence Matrices (TMCM) of second and third order [3], combined with highly performant conventional classifiers, such as the Support Vector Machines (SVM), the Multilayer Perceptron (MLP), Random Forest (RF), respectively AdaBoost in conjunction with Decision Trees (J48) were involved in this phase. Recently, taking advantage of the spectacular development of deep learning methods, various types of such techniques were considered for employment and assessment within ultrasound images for HCC recognition [4–6]. In the current approach, we focus on those deep learning techniques that led to the best classification performance in our research, also on the combinations among these techniques, as well as on their combinations with the Conventional Machine Learning (CML) methods, at both classifier and decision level.

## **2. The state of the art regarding tumor recognition in medical images**

Formerly, texture analysis methods, combined with traditional classification techniques were widely implemented for achieving automatic and computer assisted diagnosis of various affections, particularly of the tumor structures, based on medical images [7–10]. A representative approach was presented in [10], the purpose being that of differentiating, from contrast enhanced CT images, between pancreatitis and pancreatic cancer (ductal pancreatic adenocarcinoma), employing textural parameters and multivariate logistic regression. In this process, the authors involved textural features as those derived from the Run-Length Matrix, respectively from the GLCM matrix, and also the sum of the voxel values. The role of multivariate logistic regression was also that of identifying relevant CT images, respectively relevant textural attributes. The textural features were calculated during the arterial and portal phases, by employing specific software tools (AnalysisKit). The method was evaluated by employing the Area under the Receiver Operating Characteristic (AuC) metric, which presented values between 84% and 98% for different groups of attributes.

During the last years, the deep learning methods demonstrated their efficiency, overpassing the performances of the conventional techniques, and being successfully employed for medical image recognition. The main aspects concerning the state of the art regarding the deep learning techniques are highlighted below.

### **2.1 Deep learning methods for medical image recognition**

A CNN having as objective the classification of the HCC tumors was presented in [11]. This network combined parallel convolutions with atrous pooling modules, for

performing HCC recognition within ultrasound images. In [12] the authors present a multi-scale neural network that classifies liver regions by employing VGG16 and InceptionV4 type CNNs. Regarding the approach presented in [13], the authors describe a method for recognizing pancreatic tumors based on contrast enhanced CT images. A multimodal network was proposed, comprising a pyramid of augmented features, a feature fusion module, respectively a dependency computation module. At the end, an AuC of 0.9455 resulted. In [14] the authors present a method based on CNNs for classifying prostate cancer of type T1 and T2 based on MRI images. An automated system for prostate biopsy classification was presented in [15], a deep learning architecture called CarconNet being involved. This network assumed the participation of ResNet50, followed by a Fully Connected Network (FCN). This approach also considered tumor evolution stage detection through the Gleason score, the most powerful prognostic predictor for patients suffering from prostate cancer.

## **2.2 Combination between conventional and deep learning techniques for medical image recognition**

Relevant approaches in this area were described in [16–21]. A representative approach was depicted in [16], where the authors fused deep learning features with classical radiomic features to predict lung malignant nodules within tomography images. The feature maps yielded by VGG-type CNNs, as well as by original CNNs, were combined with classical radiomic features, which included shape, size, GLCM, Wavelet and Laws features. A Symmetric Uncertainty (S.U.) method was adopted for attribute selection, from the deep learning, respectively from the set of traditional features, then the combined feature vector was provided to a classifier of type RF. The most increased accuracy, 76.79%, was achieved when adopting a VGG-type CNN. In [17] the authors combined handcrafted and deep learning features for achieving automatic prostate cancer diagnosis from transrectal ultrasound images. The handcrafted features comprised robust and scale-invariant features, features referring to orientation, shape, texture, and color. The deep learning features were derived from an original CNN, consisting of a backbone and another network, as well. Different backbones were assessed, such as ResNet18, ResNet50, VGG11, VGG16, respectively DenseNet121 and DenseNet201. The concatenated feature vector was fed at the input of a fully connected network that completed the fusion process. The best classification accuracy, of 95.54%, the maximum specificity value, of 93.64%, the most increased value for the sensitivity, of 97.27%, respectively the maximum AuC, of 98.24%, resulted in the case when ResNet18 was adopted as the backbone. On the other side, assessing the handcrafted and deep learning features separately, provided a classification accuracy below 90%. Other approaches fused multiple types of deep learning features [18, 19]. A representative approach aiming for breast tumor classification was described in [18], where the authors trained deep CNNs of type GoogleNet and AlexNet. The deep learning features extracted from these structures were combined, after feature selection, with textural features. The fused feature vector was finally fed to a SVM classifier. In [19], the authors assessed the fusion between the deep learning vectors yielded by the ResNet50 and DenseNet201 CNNs, to perform automatic recognition of brain tumors. After a feature selection phase, the resulting feature vectors were combined through a serial procedure, respectively fed to a SVM classifier, yielding a recognition rate of 87.8%. A common approach with regards to the fusion between classifiers is that of considering them separately, followed by the application of a voting procedure upon the output vectors [20, 21]. Soft (average)



voting implies performing an arithmetic mean among the output vectors, between the corresponding class probabilities. Hard voting means majority voting, multiple classifiers being involved, the most frequently detected class constituting the final class. Adaptive voting implies that an important weight is provided through a separate classifier, being assigned to each model thereafter [20]. An adaptive voting-based technology was presented in [21], for achieving lung cancer classification. It has been supposed that not all the classifiers brought the same contribution to the recognition result, so the contribution weights were learnt through a specific algorithm. Thereafter, a 3-D CNN classifier, besides conventional classifiers, as for example SVM or logistic regression, were trained to predict the lung cancer type (benign or malignant), based on low-dose chest CT images. The corresponding output probabilities were exploited to train a second-stage classifier that yielded the result. At the end, a maximum AUC of 75% was achieved, due to the ensemble classifiers.

As we notice from the above-presented methods, various types of techniques were involved for tumor recognition within medical images. However, the classification performance can be further improved, for achieving a reliable automatic and computer aided diagnosis. Also, the problem of HCC recognition within ultrasound images was less approached and no systematic analysis exists upon various classes of methods applied for this purpose. In the next sections, we present our own contribution regarding each class of techniques (CML methods, deep learning techniques and combinations).

### **3. Comparing and combining conventional and deep learning techniques for HCC recognition within ultrasound images**



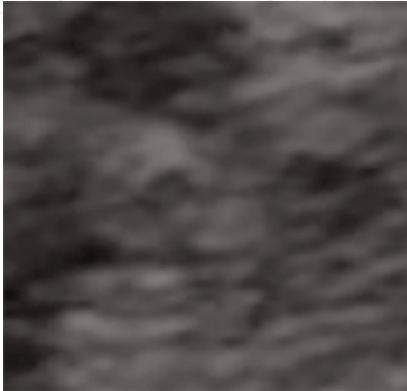

#### **3.1 The dataset and the experimental settings**

**The dataset:** For achieving reliable results, two datasets were involved in our experiments. The first one, GE7, comprised classical (B-mode) ultrasound images, corresponding to 200 cases of HCC, acquired through a Logiq 7 (General Electric, USA) ultrasound machine, considering the same acquisition parameters: 5.5 MHz frequency, gain having the value of 78, depth 16 cm, Dynamic Range (DR) having the value 111. The second dataset, GE9, included B-mode ultrasound images that belonged to 96 patients suffering from HCC, acquired through a Logiq 9 (General Electric, USA) ultrasound machine, the acquisition parameters being as follows: 6.0 MHz frequency, gain having the value 58, depth having the value of 16 cm, DR being 69. All the images were collected by the specialists from well-known clinics in Cluj-Napoca: from the 3rd Medical Clinic, respectively from the O. Fodor Regional Institute of Gastroenterology and Hepatology. All the patients in our study were subjected to biopsy or CT exam, for diagnostic confirmation. For each patient, multiple images were included. Two classes were considered for differentiation in our study: HCC and the cirrhotic parenchyma on which HCC had evolved (PAR). For the GE7 dataset, rectangular regions of interest, having the dimension of 50x50 pixels, were selected through manual procedures, by the specialists, inside HCC, or within the PAR region, employing a specific application. The second dataset comprised advanced phase HCC, these tumors being manually delineated by the physicians, using the VIA tool [22]. Through the graphical interface of this application, the doctors delimited the HCC area by using a polygon. Thus, rectangular regions of interest (patches), having  $56 \times 56$  pixels in dimension, were automatically provided from the HCC and PAR

areas, through a sliding window algorithm [4]. Representative patch examples for each of the GE7 and GE9 datasets, from the HCC and PAR classes, are provided in **Table 1**. For both datasets, we remark the heterogeneous character of HCC, as compared with PAR.

**Experimental settings:** Most of the CNNs, i.e. ResNet101, InceptionV3, Efficientnetb0, respectively the improved versions of Efficientnetb0, were employed in Matlab R2021-b, using the Deep Learning Toolbox [23]. The enhanced Efficientnetb0, improved with the aid of an ASPP module, respectively of a dropout layer, was constructed in the context of the Deep Network Designer. The training of these networks was performed as follows: the Stochastic Gradient Descent with Momentum (SGDM) strategy; the learning rate was 0.0002; the mini-batch size was 30; the momentum was considered 0.9; the duration of training was 100 epochs.

These hyper-parameters were set in this manner for achieving an accurate, efficient learning process, respectively for simultaneously avoiding overtraining, also considering the memory constraints of the computer (the minibatch size). All the networks mentioned above were pretrained on the ImageNet dataset. The ConvNext type CNN was employed in Python, by using the TorchVision library [24]. This network was trained similarly, with the same strategy and the same hyper-parameter values as those adopted for the other CNNs. As it concerns the dimensionality reduction methods, the KPCA technique [25] was implemented in Matlab 2021, by using the Matlab-Kernel-PCA toolbox [26], considering the linear, third-degree polynomial and Gaussian kernels. The Particle Swarm Optimization (PSO) [27] based feature selection technique was employed in Matlab as well, using a specific framework [28]. The

Image set	HCC	PAR
GE7		
GE9		

**Table 1.**  
 Relevant examples of HCC and PAR patches from the GE7 and GE9 datasets.

classical feature selection techniques were employed by using Weka 3.8. [29]. The CfsSubsetEval method, standing for Correlation based Feature Selection (CFS) [30], was implemented with BestFirst search, while the Information Gain Attribute Evaluation (IGA) technique was employed with Ranker search [29, 30]. The conventional classifiers were also employed using Weka 3.8 [29]. The John Platt's Sequential Minimal Optimization (SMO) methodology [29], which is equivalent with SVM in Weka, was assessed, in combination with the polynomial kernel, of 3rd degree that yielded the best performance. The metaclassifier of AdaBoost was evaluated for 100 iterations, together with the J48 method, the Weka equivalent of the C4.5 algorithm. The RandomForest classification technique was also adopted in Weka 3.8. Part of the textural features were derived with the aid of our own Visual C++ modules, as presented in Section 2.3, in a manner independent on orientation, scale and illumination, after the application of a median filter for speckle noise attenuation. The LBP feature vector was calculated in Python, with the aid of the Numpy library. All these experiments were conducted on a computer with a 2.60 GHz i7 processor, 8 GB of internal (RAM) memory, respectively an Nvidia Geforce GTX 1650 Ti GPU. As for the performance assessment strategy, in the case of the CNN based techniques, 75% of the data stood for the training set, 8% of the data constituted the validation set and 17% of the data was comprised in the test set. Regarding the conventional classifiers, 75% of the data was included in the training set, and 25% of the data being integrated in the test set. The performance of the best classifiers and fusion schemes was reassessed as well by employing cross-validation with 5 folds.

**Performance assessment:** For classification performance evaluation, the following metrics, appropriate for automatic diagnosis in the medical domain were approached: recognition rate (accuracy), TP Rate (sensitivity), TN Rate (specificity), respectively Area under ROC (AuC) [31]. In the current experimental context, HCC was considered the positive class, PAR being considered the negative class.

## 3.2 Comparing conventional and deep learning techniques

### 3.2.1 Conventional techniques involved in our research

Regarding the texture-based methods, both relevant classical techniques, and advanced, original methods, developed by the authors, were considered. As classical textural attributes, we included:

- first order statistics of the gray levels (arithmetic mean, maximum and minimum values)
- second order gray level features: the Haralick parameters derived from GLCM computed as described in [4], this feature group comprising the GLCM entropy, the GLCM energy, the GLCM homogeneity, the GLCM correlation, the GLCM variance and the GLCM contrast, which emphasized the properties of the tissue, such as the heterogeneity, echogenicity, granularity, complexity of the gray level structures.
- the autocorrelation index [30], referring to the granularity of the tissue;
- edge based statistics, such as the edge frequency and edge contrast [32], emphasizing the complexity on the gray level distribution;

- the statistics of the microstructures, resulted after the application of the Laws filters [33], emphasizing the structural complexity;
- the Hurst fractal index, highlighting the roughness and structural complexity of the tissue;
- multiresolution features, such as the Shannon entropy derived after applying the Wavelet transform recursively, twice [2];
- LBP, as a powerful texture-based technique, insensitive to illumination changes [33]. For calculating these features, around each pixel, a circle of radius R was drawn, and, on this circle, N neighbors were emphasized. For achieving the LBP code, the differences between the central pixel and each neighbor were determined. For each considered neighbor, if this difference was larger than 0, a code having the value of 1 was stored, otherwise, a 0 valued code was considered. These N codes formed a number that represented the LBP code. In the current work, the LBP attributes were calculated by ranging the values of R and N, the following (R, N) value pairs being assessed: (1, 8), (2, 16), (3, 24). Compressed LBP histograms with a smaller number of bins (100), were computed in the area of interest.

We also employed advanced, original textural attributes, conceived by the authors, as for example the edge orientation variability [5], or those derived from superior order GCMs, defined by (1) and computed as described in [5].

$$C_D(a_1, a_2, \dots, a_n) = \# \left\{ ((x_1, y_1)(x_2, y_2)(x_3, y_3) \dots (x_n, y_n)) : \right. \quad (1)$$

$$A(x_1, y_1) = a_1, A(x_2, y_2) = a_2, \dots,$$

$$A(x_n, y_n) = a_n, |x_2 - x_1| = |\vec{d}x_1|, |x_3 - x_1| = |\vec{d}x_2|, \dots,$$

$$|x_n - x_1| = |\vec{d}x_{n-1}|, |y_2 - y_1| = |\vec{d}y_1|, |y_3 - y_1| = |\vec{d}y_2|, \dots,$$

$$|y_n - y_1| = |\vec{d}y_{n-1}|, \operatorname{sgn}((x_2 - x_1)(y_2 - y_1)) = \operatorname{sgn}(\vec{d}x_1 \cdot \vec{d}y_1),$$

$$\left. \operatorname{sgn}((x_n - x_1)(y_n - y_1)) = \operatorname{sgn}(\vec{d}x_{n-1} \cdot \vec{d}y_{n-1}) \right\}$$

According to (1), each element of this matrix,  $C_D(a_1, a_2, \dots, a_n)$  contains the number of n-tuples of pixels, having the values  $(a_1, a_2, \dots, a_n)$  for the considered attribute  $A$ , which can stand for the intensity level, edge orientation, etc. The pixels are in a spatial relationship denoted by the displacement vectors,  $(d_{x1}, d_{y1}), (d_{x2}, d_{y2}), \dots, (d_{xn-1}, d_{yn-1})$ . In our approach, the attribute  $A$  stood either for the gray level value of the pixel in the original image, respectively for the gray level value associated to the cluster center, after applying the k-means clustering algorithm, each resulting cluster being associated with a textural microstructure. In the case of k-means clustering, the value of  $k$  was 250 or 500. The second and third order GLCM, respectively the second and third order Textural Microstructure Cooccurrence Matrix (TMCM) were determined.



### 3.2.2 Convolutional neural networks (CNNs)

As for the deep learning techniques, we included both relevant and newly developed CNNs, considering both classical CNNs and transformer-based methods [34–38]. Several standard architectures were initially analyzed, the best performance being achieved for InceptionV3 [35], ResNet101 [36], DenseNet201 [37], respectively for the recently developed EfficientNetb0 [34]. Concerning the transformer-based methods, the best performance resulted for ConvnextBase [38]. These CNN architectures were considered as embedding both inception modules and residual connections, being well-known that these elements considerably enhanced the classification performance. The high performant densely connected networks, in the form of the DenseNet architecture, were included as well, while the EfficientNet architecture was also employed, due to its' scaling properties. Some of these architectures were enhanced, for optimizing their performances. Thus, two improved versions of Efficientnetb0, denoted Efficientnet\_ASPP1, respectively Efficientnet\_ASPP2, were conceived, by introducing, before the fully connected layer, an AtrousSpatial Pyramid Pooling (ASPP) module [34], designed in two manners, for extracting multi-scale features. A dropout layer was added thereafter, to avoid overfitting. The ASPP modules were inserted after the usual convolutional part of Efficientnetb0, immediately before the fully connected layers. The first ASPP module included a  $1 \times 1$  convolution, as well as two atrous convolutions of size  $3 \times 3$ , with the rate 2, respectively 3 (Efficientnet\_ASPP1). The second ASPP module comprised a  $1 \times 1$  convolution unit, a  $3 \times 3$  atrous convolution unit with rate 3, respectively a  $5 \times 5$  atrous convolution unit, with rate 2 (Efficientnet\_ASPP2). At the end, a *depthcat* layer, respectively a global average pooling layer were added in both cases. Regarding the dropout layer, an output probability of 0.5 was associated to it.

### 3.2.3 Experimental results

#### 3.2.3.1 Assessing the performance of the textural features through conventional classifiers

Firstly, upon the entire feature vector, which contained the above-mentioned textural features, the CFS technique in combination with the IGA method was employed for relevant feature selection, the intersection between the two resulting relevant feature subsets constituting the final feature set. Then, the values of these features were fed at the entrances of conventional classifiers, such as SVM (SMO in Weka 3.8), RF, respectively AdaBoost combined with Decision Trees (J48) [29]. For the first dataset, GE7, the best classification accuracy, of 92.92%, the most increased sensitivity, of 94.1% as well as the highest AUC, of 93.45%, were achieved for AdaBoost, while the most increased specificity, of 93.1%, resulted for SVM. As for the GE9 dataset, the highest accuracy, of 82.5%, the most increased sensitivity, of 81.81%, the maximum specificity, of 83.2%, as well as the best AUC, of 89.7%, resulted for AdaBoost as well.

#### 3.2.3.2 Assessing the performance of the convolutional neural networks (CNNs)

Within **Table 2**, the classification performance parameters for the individual CNNs, obtained through transfer learning, on both datasets, GE7 and GE9, are depicted. The highest values resulted for each classification performance parameter, for each dataset, were emphasized with bold. For GE7, the maximum classification

Dataset	Method	Accuracy	Sensitivity	Specificity	AuC
GE7	ResNet101	<b>95.9%</b>	<b>95.6%</b>	<b>91.2%</b>	<b>93.4%</b>
	DenseNet201	93.1%	92.8%	92.5%	93%
	InceptionV3	88.7%	88.8%	88.6%	89%
	Efficientnetb0	74.93%	72.9%	77.5%	75.2%
	Efficientnet_ASPP1	76.9%	77.4%	76.1%	76.75%
	Efficientnet_ASPP2	73.2%	72.1%	79.8%	77%
	ConvnextBase	83%	78%	88%	83%
GE9	ResNet101	78.4%	82%	75.5%	78.75%
	DenseNet201	<b>83.2%</b>	<b>83.5%</b>	82.9%	83.20%
	InceptionV3	80.39%	81.63%	79%	<b>86%</b>
	Efficientnetb0	74.32%	75.22%	73.22%	82%
	Efficientnet_ASPP1	76.2%	79.8%	73.22%	76.51%
	Efficientnet_ASPP2	67.1%	48.5%	<b>86.9%</b>	67.70%
	ConvnextBase	81%	75%	86%	80.5%

**Table 2.**  
 The assessment of the individual CNN architectures.

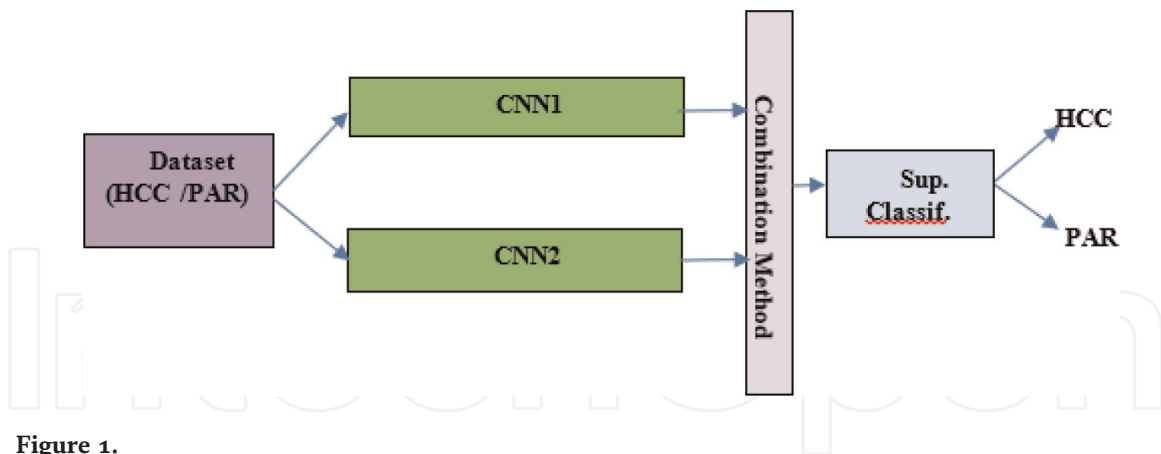
accuracy, the most increased sensitivity and the best AUC, resulted for ResNet101. The maximum specificity was achieved for DenseNet201. For the second dataset, GE9, DenseNet201 yielded the best accuracy of 83.2%, respectively sensitivity, of 83.5%. The best specificity, of 86.9%, was provided by Efficientnet\_ASPP2, while the best AuC, of 86%, resulted for InceptionV3. For both GE7 and GE9 datasets, Efficientnet\_ASPP1, the first enhanced version of Efficientnetb0, led to an increased classification performance, in terms of accuracy and sensitivity, as compared with the original Efficientnetb0, while the second improved version, Efficientnet\_ASPP2, led to an improved specificity, in comparison with the initial version, Efficientnetb0.

### 3.3 Combinations among convolutional neural networks (CNNs)

#### 3.3.1 Classifier level fusion

In this approach, we considered representative CNN architectures, then we combined them at the classifier level. The selected architectures were ResNet101, InceptionV3, respectively Efficientnetb0, these CNNs being acknowledged for their performances, as previously described. Two CNNs were combined, by extracting the features at the end of the convolutional part of the network, before the fully connected or *softmax* layers, as depicted by **Figure 1**, fusing them through the following procedures thereafter:

1. *Concatenation*, assuming the simple concatenation of the feature vectors.
2. *CFS + Concatenation*, involving feature selection through the CFS technique [28], on each feature vector, followed by concatenation.



**Figure 1.**  
Classifier-level fusion of two CNNs.

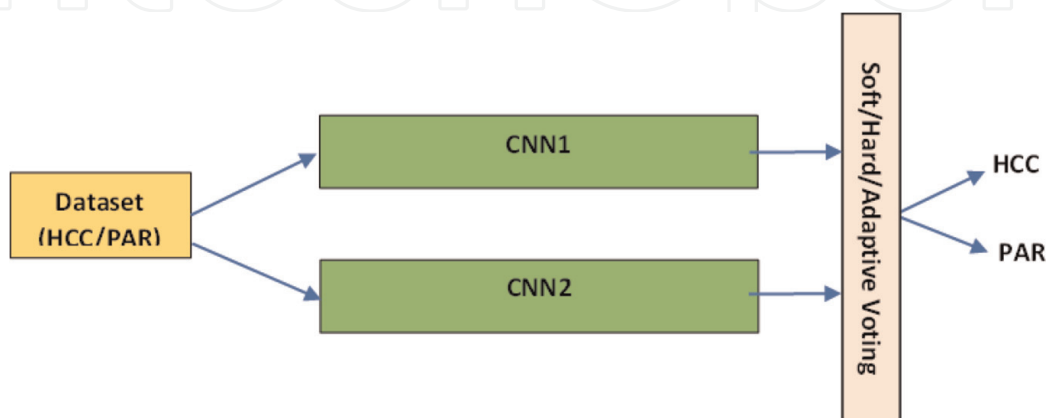
3. *Concatenation + CFS*, implying to perform the concatenation of the two vectors followed by feature selection through CFS [28].
4. *KPCA + Concatenation* assumed the application of KPCA [23] on each vector, performing concatenation thereafter.
5. *Concatenation + KPCA*, employing the concatenation of the two vectors followed by KPCA [23] .

For KPCA, the Gaussian kernel was employed, which provided the best results. The combined feature vector was fed at the input of a conventional classifier, such as SVM, RF, AdaBoost combined with the C4.5. method of Decision Trees [29].

### 3.3.2 Decision level fusion

Considering the CNN architectures previously described, we performed decision level fusion among them, implying to combine the output probabilities of these CNNs through soft, hard and adaptive voting [18], as also illustrated in **Figure 2**.

We provide below the detailed description of each procedure. Soft(average) voting: For soft voting, relevant pairs, or CNN groups, were considered, respectively an arithmetic or weighted mean was determined among the output probabilities [6]. As for the weighted mean, a larger weight was attributed to those networks having



**Figure 2.**  
Decision level fusion (voting) of two CNNs.

increased performance, according to (2). In (2),  $W_{\text{mean}}$  stood for the weighted mean,  $P_1, P_2, \dots, P_{n-1}$  constituted the predictions (probability pairs) of the classifiers in the group,  $P_n$  constituted the prediction of the best classifier, while  $P_1$  represented the prediction corresponding to the weakest classifier.

$$W_{\text{mean}} = \frac{2(n-1)P_n + 2(n-2)P_{n-1} + \dots + 2P_2 + P_1}{2(n-1) + 2(n-2) + \dots + 1} \quad (2)$$

Hard (majority) voting: Hard (majority) voting was also implemented, upon representative CNN groups. The final derived class was the most frequently met class at the output of the considered classifiers (the majority class). Adaptive voting: Adaptive voting was implemented as well, through a stacking combination scheme, by providing the CNNs' outputs (pairs of probabilities) to a conventional classifier, which yielded the final class. The following conventional classifiers were assessed: Multilayer Perceptron (MLP), Random Forest (RF), Support Vector Machines (SVM), AdaBoost in conjunction with decision trees (C4.5) [22].

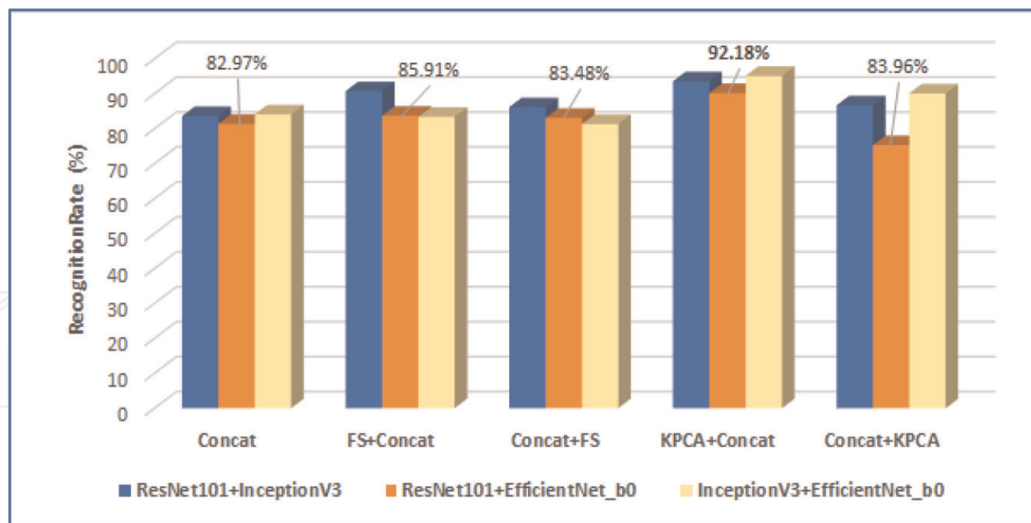
### 3.3.3 Experimental results

#### 3.3.3.1 Classifier level fusion

Regarding the assessment of the CNN combinations at classifier level, the CNN features were firstly extracted: 2048 features were obtained from ResNet101, respectively InceptionV3, at the output of the *pool5*, respectively *avg\_pool* layers, while for Efficientnetb0, 1280 features were extracted at the end of *GlobAvgPool*. On both datasets, GE7 and GE9.

**Performance assessment on the GE9 dataset:** The highest accuracy, of 97.79% resulted when considering the fusion between ResNet101 and InceptionV3, for the *KPCA + Concatenation* combination scheme, when considering the AdaBoost metaclassifier. The most increased sensitivity, of 97.9%, was achieved in the case of the classifier level fusion between ResNet101 and InceptionV3, for *KPCA + Concatenation*, for the RF classifier, respectively in the case of the combination between ResNet101 and Efficientnetb0, for the same combination scheme and classifier as previously. The best specificity, of 98.9%, respectively the best AuC, of 97.8% were attained when fusing ResNet101 and InceptionV3, through the *KPCA + Concatenation* scheme for the AdaBoost metaclassifier. The best classification performance usually resulted for *KPCA + Concatenation*, respectively *CFS + Concatenation*, when employing the RF, respectively the AdaBoost metaclassifiers. In **Figure 3**, a graphical representation of the average accuracy, for each fusion scheme, for each CNN combination, is provided. Above, the arithmetic mean of the classification accuracies, for each fusion scheme, considering all the CNN combinations, is depicted: the best performance was achieved for *KPCA + Concatenation*, the overall average accuracy being 92.18%, followed by the *FS + Concatenation*, the overall average accuracy being 85.91%. We observe that performing first feature selection or KPCA, on each CNN feature vector, separately, followed by concatenation, led to better results than the procedures which performed firstly concatenation, then FS or KPCA. We also notice that the simple concatenation of the CNN features was overpassed by all the other fusion schemes. Regarding the CNN combinations, the best results were provided by *ResNet101 + InceptionV3*, followed by *InceptionV3 + Efficientnetb0*, respectively by *ResNet101 + Efficientnetb0*.





**Figure 3.** Comparison of the average classification accuracies for each fusion scheme, for every CNN combination.

The classification performance of the best fusion schemes also exceeded the performance of the individual CNNs (Table 2).

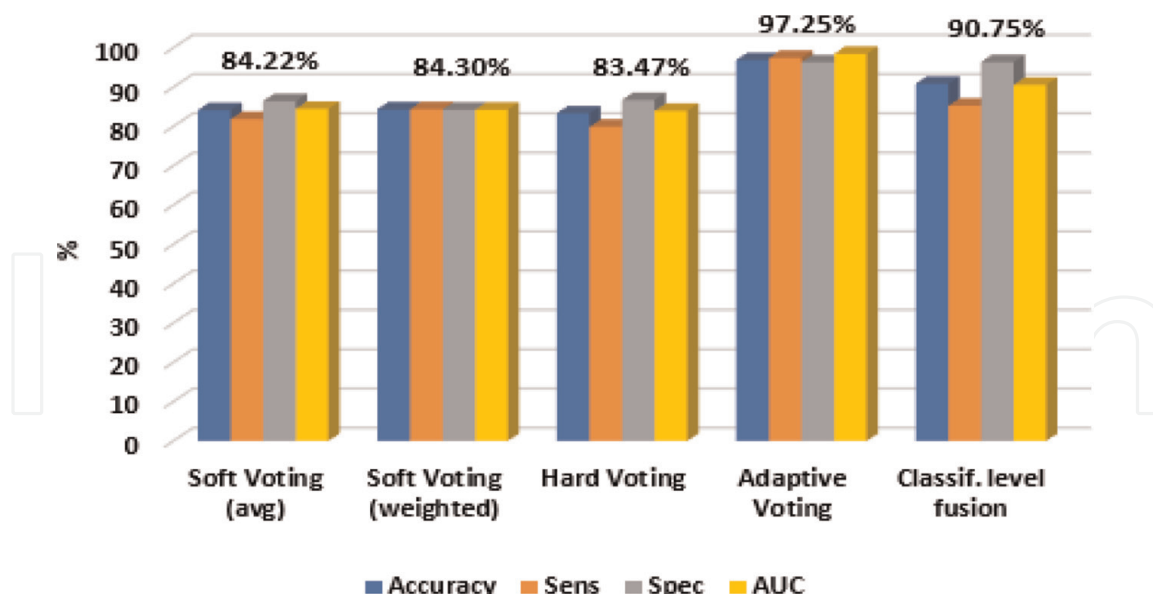
**Performance assessment on the GE7 dataset:** For assessing the performance of the CNN combinations at classifier level on the GE7 dataset, the same pairs of CNNs as in the previous case were assessed. The best classification performance was attained for ResNet101 combined with InceptionV3, for the KPCA+Concatenation scheme: accuracy of 98.71%, sensitivity of 99.3%, specificity of 98.1%, AuC of 99.8%.

### 3.3.3.2 Decision level fusion

In the current study, the experiments regarding decision-level fusion of the individual CNNs were performed on the GE9 dataset, more recently acquired.

When employing *soft voting*, through the arithmetic mean between the output probabilities of the CNNs, pairwise combinations, as well as CNN groups were considered. When applying *hard(majority) voting*, three relevant groups of CNN were considered: the top three CNNs with the best classification performance, ResNet101, DenseNet201 and InceptionV3; the fusion of the two best performing CNNs with Efficientnet\_ASPP1 only, as this enhanced EfficientNet version led to a more increased performance than Efficientnet\_ASPP2; the fusion of the top three CNNs, having the maximum performance, with Efficientnet\_ASPP1, respectively Efficientnet\_ASPP2.

In Figure 4, a comparison of the classification performances for all the fusion schemes is illustrated. The arithmetic means of the performance metrics, calculated for each scheme, were considered. Classifier level fusion was also evaluated, on the same dataset. Above each group that corresponded to a certain fusion scheme, the arithmetic mean of the performance parameters is emphasized. We can observe that adaptive voting led to the most increased performance, the average accuracy being 96.85%, the average sensitivity being 97.43%, the average specificity being 96.30%, respectively the average AuC being 98.45%. On the second place we can find classifier level fusion, the average accuracy being 90.83%, the mean sensitivity being 85.27%, mean specificity being 96.34%, the mean AuC being 90.59%. Regarding the decision



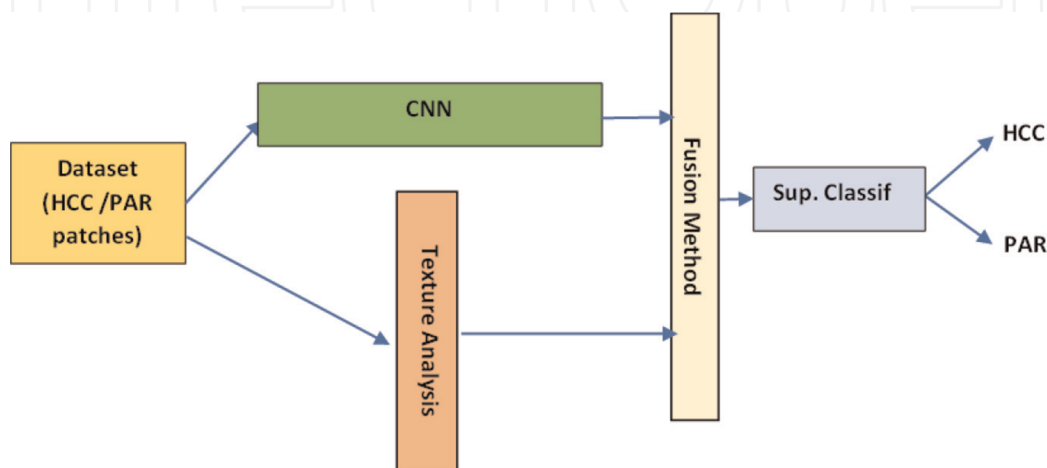
**Figure 4.**  
 Classification performance comparison between the CNN based fusion schemes.

level fusion schemes, adaptive voting generated the best performance, followed by soft voting, respectively by hard voting.

### 3.4 Combinations between conventional techniques and convolutional neural networks (CNNs)

#### 3.4.1 Classifier level fusion

The classifier level combination of the CNN based techniques with the CML methods assumed to feed the initial dataset, consisting of HCC and PAR patches, as input to a CNN classifier, respectively to the texture analysis techniques, as illustrated in **Figure 5**. The deep learning attributes, resulted at the termination of the of CNN convolutional part, were combined with the textural attributes, through concatenation, or through a fusion scheme that involved dimensionality reduction, as KPCA or Feature Selection (FS). Then, a supervised conventional classifier was employed. The textural features presented previously formed the vector of conventional features,



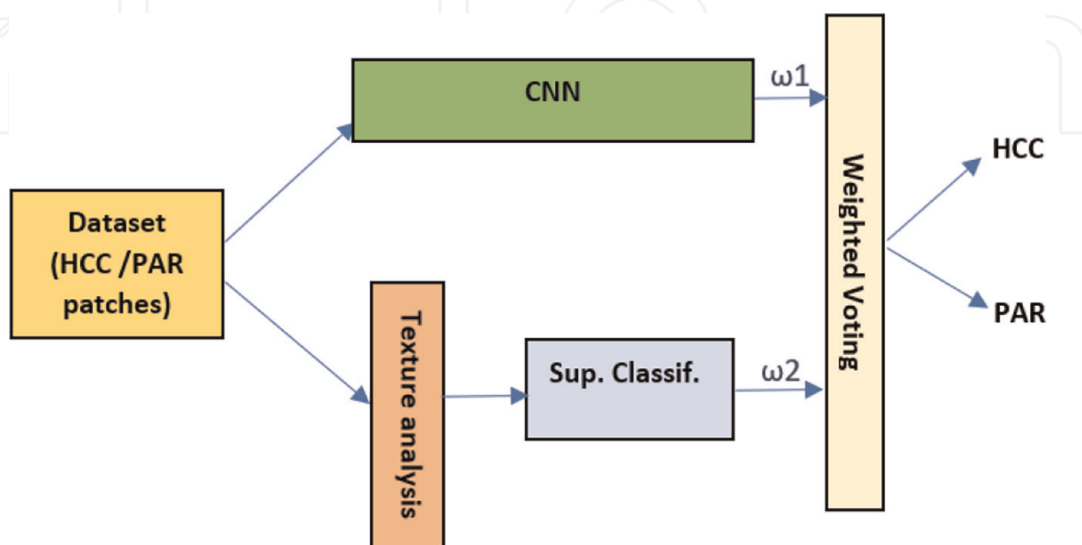
**Figure 5.**  
 Combination of the conventional and deep learning techniques at classifier level.

while the deep learning-based features were collected at the end of the last layer that preceded the fully connected layers. The fusion methods were employed through the same combination schemes as mentioned in Section 3.3.1. However, in this situation, feature selection has been performed by employing both classical and bio-inspired techniques. In the case of classical feature selection, the CFS and IGA methods were adopted. The CFS technique was employed in conjunction with the *best first search* algorithm, while IGA was implemented together with the *Ranker* search algorithm. These two FS methods were combined, by finally providing the intersection between the two resulting feature subsets. Regarding the bio-inspired techniques, the Particle Swarm Optimization (PSO) algorithm was implemented as described in [3]. Thus, the corresponding Fitness function aimed to minimize the classifier error rate, but also the number of relevant features included in the final feature set.

Regarding KPCA, considering *Concatenation + KPCA*, we retained 500 components, while for *KPCA + Concatenation*, we retained 300 components from each vector, containing either deep learning or textural features, to balance the vector lengths in each case. Thereafter, the correlations between the deep learning and textural features were analyzed, to explain the importance of the deep learning attributes reported to the physical and visual properties of the tumor. To do so, the Pearson correlation technique was adopted [5].

### 3.4.2 Decision level fusion

The fusion between the CNN based techniques and the CML methods at the decision level, implies that the CNN classifier, respectively the CML method, consisting of advanced texture analysis and a conventional supervised classifier, are firstly applied in parallel, followed by the combination of the output probabilities of the two classifiers (deep learning and traditional), as illustrated in **Figure 6**. In the current work, the weighted mean operation was performed between the probability output vectors. Thus, the final output of the combined classifier, the Decision Level output (*DL\_output*), consisting of the probability vector indicating the class probabilities, was derived by employing the formula (3). In (3),  $\omega_1$  and  $\omega_2$  are the weights associated with the two terms of the weighted mean, the method yielding the best



**Figure 6.** Combination of the conventional and deep learning techniques at the decision level.

classification accuracy being assigned a weight of value 2, while the weight associated with the other method will have the value 1.  $P_1$  was the output (pairs of probabilities) of the CNN, while  $P_2$  was the output (pairs of probabilities) of the conventional supervised classifier.

$$DL_{\text{Output}} = (\omega_1 * P_1 * + \omega_2 * P_2) / (\omega_1 + \omega_2) \quad (3)$$

In the current approach, the best performing CNN was fused with the best CML method, assuming the relevant textural features and the conventional classifier yielding the highest accuracy.

### 3.4.3 Experimental results

#### 3.4.3.1 Classifier level combination

##### 3.4.3.1.1 Performance assessment on the GE7 dataset

For assessing and comparing the considered fusion schemes, we computed the arithmetic mean of the values of the performance parameters, attained for the individual traditional classifiers, for each fusion between the textural features and the deep learning features obtained from a certain CNN.

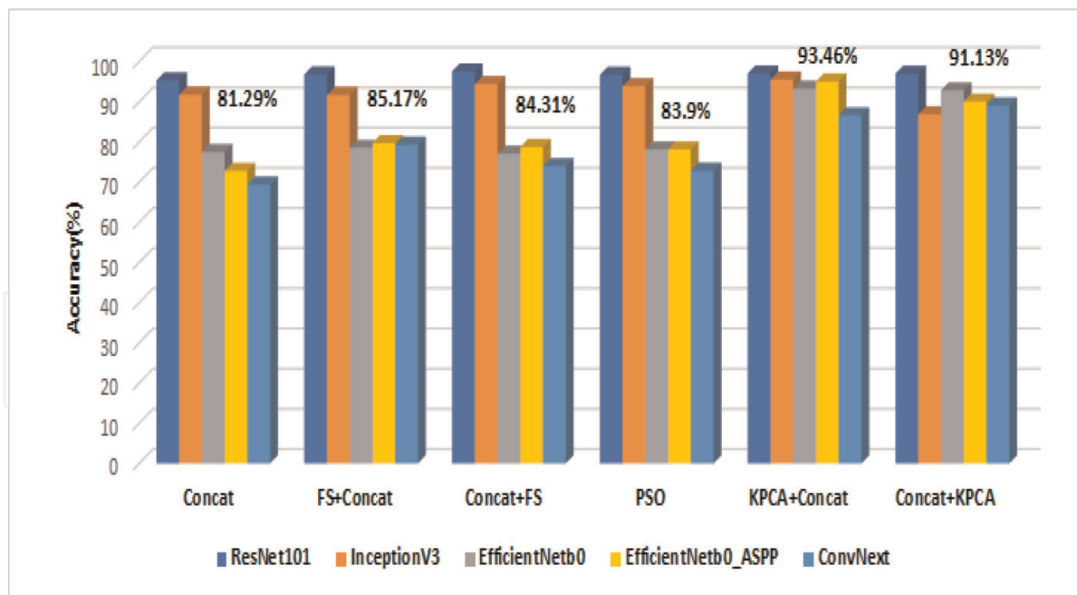
The overall maximum value of the mean accuracy, 97.47%, as well as the overall maximum value of the mean sensitivity, of 97.53%, were achieved when combining ResNet101 with the textural features, through the *Concatenation + FS* combination scheme, the overall highest value of the average specificity, of 98.63%, was attained when fusing the InceptionV3 feature vector with the textural feature vector, for the *KPCA + Concatenation* combination, while the overall highest value of mean AUC, of 97.86%, was achieved when fusing ResNet101 with the textural features through the *Concatenation + PSO* scheme. The highest overall accuracy, 98.23%, was attained when InceptionV3 was considered, for *KPCA + Concatenation*, in the case of AdaBoost. The maximum overall sensitivity, 98.2%, was achieved when ResNet101 was employed, for *KPCA + Concatenation*, respectively for AdaBoost.

Above each group, corresponding to a certain fusion scheme, the arithmetic mean of the accuracy values per group was emphasized. As can be noticed, the performance of the considered combination schemes exceeded that of the individual CNNs, in most of the situations. Also, all the fusion schemes employing feature selection and KPCA, led to a better performance than that achieved when considering a simple concatenation between the CNN and the textural feature vectors. Thus, the highest value of the average accuracy, 93.46%, was attained for *KPCA + Concatenation*, followed by the average accuracy of 91.13%, achieved for the *Concatenation + KPCA* combination (Figure 7).

##### 3.4.3.1.2 Performance assessment on the GE9 dataset

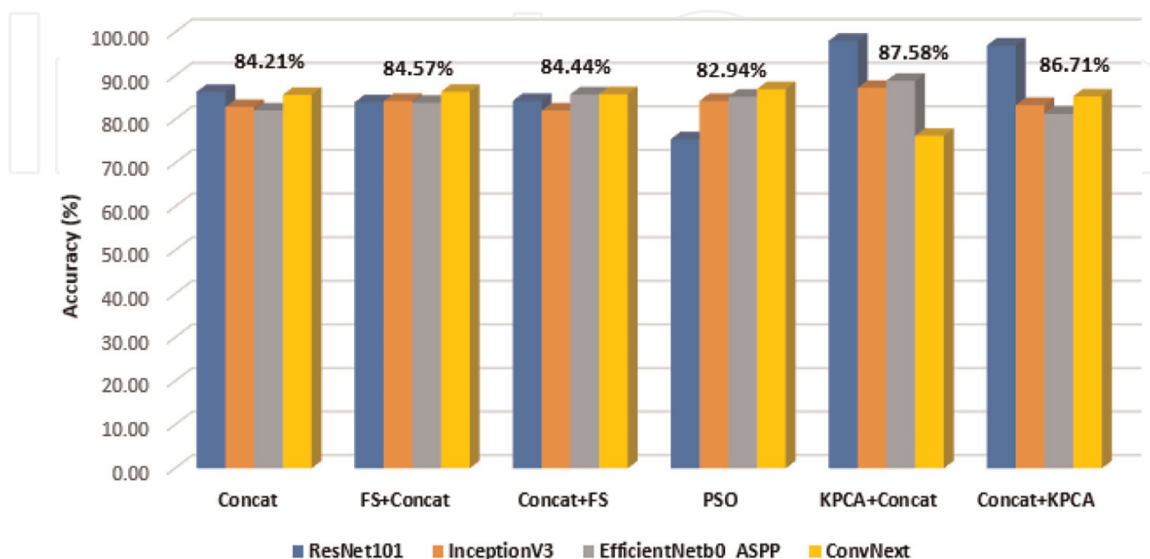
Similarly, with the case of the GE7 dataset, we first computed the arithmetic mean of the values of the performance parameters, obtained, for the three traditional classifiers, for each fusion between the textural features and the deep learning features derived from a particular CNN. The maximum value of the mean accuracy, of 98.01%, the highest mean sensitivity, of 98.26%, the most increased mean specificity, of 97.9%, respectively the maximum mean AUC, of 94.16% were achieved for the





**Figure 7.** Comparisons between the mean accuracy values resulted for each fusion scheme, for the employed CNNs, for the dataset GE7.

*KPCA + Concatenation* fusion scheme, when the ResNet101 architecture was employed. Regarding the individual values, attained through each traditional classifier, the maximum overall accuracy, 98.9% and the highest overall specificity, 98.6%, resulted for the fusion between InceptionV3 and the textural attributes, for the *KPCA + Concatenation* combination, when considering the AdaBoost metaclassifier. In **Figure 8**, the comparison between the values of the arithmetic means of the accuracies, for each combination scheme, for each CNN, is depicted, with the arithmetic average of the accuracies per combination scheme being illustrated for each group. The highest mean accuracy, 87.58%, was achieved for KPCA, followed by concatenation, while the second-best mean value, of 86.71%, was attained for concatenation followed by KPCA.

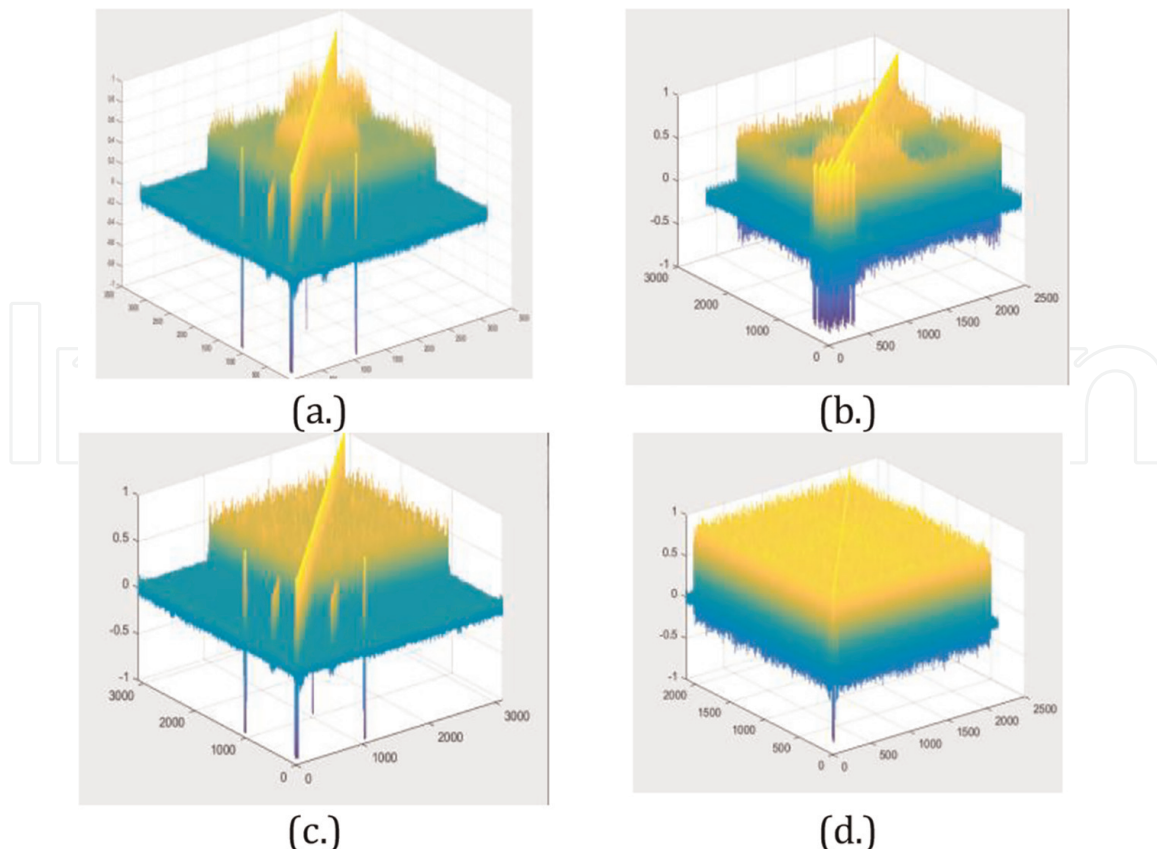


**Figure 8.** Comparisons between the mean accuracy values resulted for each fusion scheme, for the employed CNNs, for the dataset GE9.

**Relevant textural attributes:** The importance of the textural parameters with respect to the classification process was evaluated considering the whole feature vector, when fusing them with the CNN feature vector. The ordering was derived after the application of the IGA technique [28] upon the combined feature vector, containing both textural and CNN features. For GE7, the most important feature was the TMCM contrast, calculated for  $k = 500$ , with an average relevance score of 0.066, the highest average score for the whole feature set being 0.357. This attribute emphasized the more complex character of the HCC tissue. Regarding the GE9 dataset, the first position among the ordered feature vector was occupied by the homogeneity computed from TMCM of order three, the value of  $k$  being 250, having associated the maximum relevance score of 0.2. This parameter highlighted the heterogeneous nature of the HCC tissue, due to the interleave of multiple tissue types.

**Correlations among the textural and CNN features:** The correlations between the textural parameters and the CNN features were also assessed, for each CNN, on both datasets. Regarding the correlations between the textural and the CNN attributes, on GE7, the most increased correlations were those between the TMCM500\_contrast and five InceptionV3 features, the values of the corresponding correlation coefficients being 0.197, 0.194, 0.184, 0.176 and 0.171. As for the GE9 dataset, the highest correlations were those met between the GLCM\_variance and three ResNet101 features, the most increased correlation coefficients being 0.429, 0.26, respectively 0.17, followed by those between the TMCM500\_contrast and the ResNet101 features of 0.176.

**Figure 9** graphically illustrates these correlations for InceptionV3, respectively for ResNet101 in the case of both GE7 and GE9 datasets. In the case of GE7, the textural



**Figure 9.** The correlations between the CNN and textural features: (a), (b) for inception V3; (c), (d) for ResNet101; (a), (c) on the GE7 dataset; (b), (d) on the GE9 dataset.

attributes have indexes in the range 1–550, while the CNN features’ indexes are in the range 51–2598. As for GE9, the textural features have associated indexes from 1 to 98, while the CNN features have indexes from 99 to 2047.

### 3.4.3.2 Decision level combination

When employing decision-level fusion, the classification performance was tested on both GE7 and GE9 datasets, considering, in each case, the best performing CNN architecture, respectively the subset of relevant textural features in conjunction with the best conventional classifier. The values corresponding to the classification performance parameters are depicted in **Table 3**.

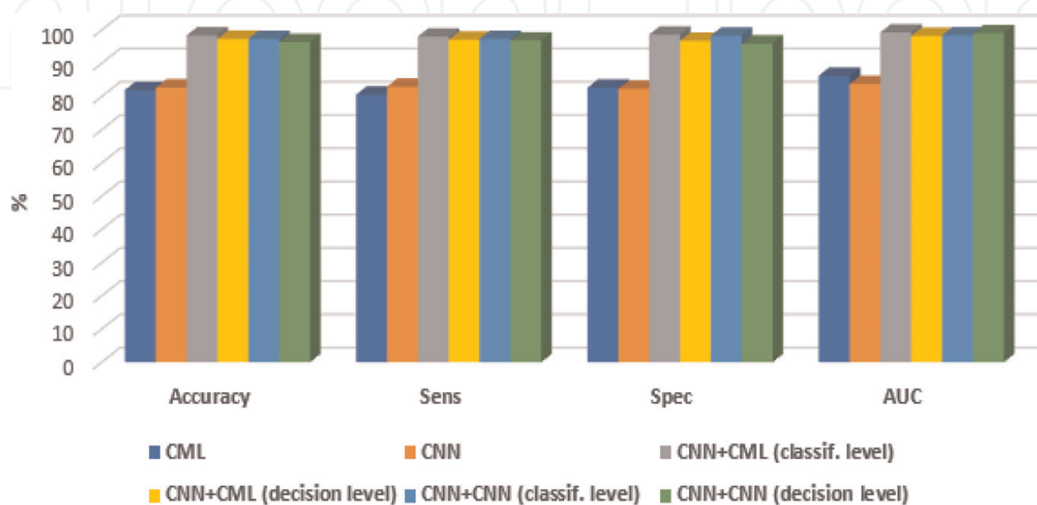
For the GE7 dataset, the weighted voting procedure was performed between the ResNet101 architecture and the AdaBoost metaclassifier, a larger weight being assigned to ResNet101, while for the GE9 dataset, the DenseNet201 architecture and the AdaBoost conventional classification technique were combined through weighted voting, the highest weight being attributed to DenseNet201. We notice that the values of the classification performance parameters corresponding to the decision level fusion between the conventional and the deep learning techniques are almost equal, however, slightly smaller than those corresponding to classifier level fusion between these classes of methods.

## 4. Discussions

In **Figure 10**, the overall comparisons between the best values of the performance parameters, for the methods involved in this study, are depicted, in the case of the

Dataset	Accuracy	Sensitivity	Specificity	AuC
GE7	97.2%	97.3%	97.7%	97.5%
GE9	97.8%	97.5%	97.3%	97.7%

**Table 3.**  
The values of the classification performance parameters for decision level fusion.



**Figure 10.**  
Comparison of the performance parameter values for all the fusion schemes.

Combination	Accuracy	Sensitivity	Specificity	AuC	Average performance
CML	81.9	80.5	82.6	86.22	82.81
CNN	82.65	82.9	82.35	83.75	82.91
CNN + CML (classifier level)	<b>98.35</b>	<b>98</b>	<b>98.6</b>	<b>99.25</b>	<b>98.55</b>
CNN + CML (decision level)	97.35	97.1	96.75	98.25	97.36
CNN + CNN (classifier level)	97.3	97.3	98.3	98.4	97.83
CNN + CNN (decision level)	96.39	96.8	95.9	99	97.02

**Table 4.** Comparison of the performance parameter values obtained for all the considered combination schemes through cross-validation.

more recently gathered dataset, GE9. For improving the reliability of the assessment procedure, the performances of the best classifiers and combination schemes resulted for each class of techniques were reassessed by employing cross-validation with 5 folds, the values of the performance parameters being also illustrated in **Table 4**. Within **Table 4**, the last column illustrates the average performance, expressed through the arithmetic mean of the performance parameter values in the case of each fusion scheme. Also, the overall maximum values for each performance parameter are highlighted in bold.

We can infer that the maximum performance resulted in the case of classifier level fusion between the conventional and deep learning techniques. It was followed by that resulted in the case of classifier level fusion between the CNN structures, then by decision level fusion between the two classes, CML and CNN, respectively by decision level fusion between the CNNs. We also notice that all the considered fusion schemes exceeded, with respect to the classification performance, both the individual CNNs, as well as the CML methods.

Regarding the comparisons with the already existing state of the art approaches, we reproduced the method presented in [16] on the experimental dataset GE9 employed in the current approach, as described in [5]. In the current approach, we reassessed the classification performance, by employing cross-validation with 5 folds. An accuracy of 84.3% resulted in this manner. The approach described in [19] was reproduced on the GE9 dataset, as well. After assessing the classification performance by employing cross-validation with 5 folds, an accuracy of 89.9% was achieved [5]. We also reproduced the approach presented in [18] on the GE9 dataset, by training an AlexNet CNN, respectively a GoogleNet CNN on the GE9 dataset. Thereafter, the deep learning features were extracted at the end of the convolutional layers, then the CFS technique was applied on each feature set to yield the relevant characteristics from each class. The resulting feature sets were combined with the relevant textural feature set, the classification accuracy being assessed through the conventional classifiers mentioned in the Section 3.4.1, by employing cross-validation with 5 folds. A detailed comparison with the state-of-the-art approaches is depicted within **Table 5**. In the last column, the difference in accuracy between our approach and each state-of-the-art approach is illustrated.

Thus, all the employed combination schemes exceeded the classification performance of these approaches, with respect to the maximum accuracy resulted in each case. As it results from **Table 5**, an average accuracy difference of 11.15%, between the current approach and the already existing state-of-the-art approaches was achieved.



Method	Accuracy	Accuracy difference
Lung nodules recognition [19]	89.9%	8.45%
Breast tumor recognition [18]	87.85%	10.5%
Brain tumor recognition [16]	84.3%	14.05%
Current approach	98.35%	—

**Table 5.**  
*Comparison with the already existing state of the art approaches.*

## 5. Conclusions and future work

The conventional and the deep learning techniques described above, together with their combinations, led to a very good classification performance, maximum classification accuracies above 95% resulting. The combination between the CML and CNN classes of techniques provided very good results, for both classifier level and decision level fusion, while applying the same fusion schemes to combine CNNs led to satisfying results, as well. In our further research, we aim to compare multiple voting procedures regarding the decision level fusion between the CMM and CNN based techniques. The current datasets based on ultrasound images will be further extended, while other types of medical images, such as CT and MRI images will be also involved for enhancing the computer aided and automatic diagnosis of HCC, performed in a non-invasive manner.

## Acknowledgements

This research was supported by the Romanian National Authority of Scientific Research and Innovation, CNCS – UEFISCDI, project number PN-III-P1-1.1-TE-2021-1293, Nr. TE 156/2022, within PNCDI III.

The work was financed as well by CLOUDUT, a project cofounded by the European Fund of Regional Development, part of the Competitiveness Operational Program, 2014-2020, contract no. 235/2020.

IntechOpen

## Author details

Delia Mitrea<sup>1\*</sup>, Sergiu Nedevschi<sup>1</sup>, Mihai Socaciu<sup>2</sup> and Radu Badea<sup>2</sup>


1 Technical University of Cluj-Napoca, Cluj-Napoca, Romania

2 Iuliu Hațieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania

\*Address all correspondence to: [delia.mitrea@cs.utcluj.ro](mailto:delia.mitrea@cs.utcluj.ro)

## IntechOpen

---

© 2023 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Elmohr K, Elsayes M, Chernyak V. LI-RADS: Review and update. *Clinical Liver Diseases*. 2021;17(3):108-112
- [2] Sherman M. Approaches to the diagnosis of hepatocellular carcinoma. *Current Gastroenterology Reports*. 2005; 7:11-18
- [3] Mitrea D et al. Automatic recognition of the hepatocellular carcinoma from ultrasound images using complex textural microstructure co-occurrence matrices (CTMCM). In: *Proceedings of 7th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*. Setubal, Portugal: Scitepress Digital Library; 2018. pp. 178-189
- [4] Brehar R, Mitrea D, Vancea F, Marita T, Nedevschi S, Lupsor-Platon M, et al. Comparison of deep-learning and conventional machine-learning methods for the automatic recognition of the hepatocellular carcinoma areas from ultrasound images. *Sensors*. 2020;20: 3085
- [5] Mitrea D, Brehar R, Nedevschi S, Lupsor-Platon M, Socaciu M, Badea R. Hepatocellular carcinoma recognition from ultrasound images using combinations of conventional and deep learning techniques. *Sensors*. 2023; 23(5):1-29
- [6] Mitrea D, Brehar R, Mocan C, Nedevschi S, Socaciu M, Badea R. Hepatocellular carcinoma recognition from ultrasound images by fusing convolutional neural networks at decision level. In: *Proceedings of the 26th International Conference on Telecommunications and Signal Processing (TSP)*. Prague, Czech Republic: IEEE; 2023
- [7] Yoshida H, Casalino D. Wavelet packet based texture analysis for differentiation between benign and malignant liver tumors in ultrasound images. *Physics in Medicine and Biology*. 2003;48:3735-3753
- [8] Sujana H, Swarnamani S. Application of artificial neural networks for the classification of liver lesions by texture parameters. *Ultrasound in Medicine & Biology*. 1996;22:1177-1181
- [9] Duda D, Kretowski M, Bezy-Vendling J. Computer aided diagnosis of liver tumors based on multi-image texture analysis of contrast-enhanced CT. Selection of the most appropriate texture features. *Studies in Logic, Grammar and Rhetoric*. 2013;35:49-70
- [10] Ren S, Zhang J, Chen J, et al. Evaluation of texture analysis for the differential diagnosis of mass-forming pancreatitis from pancreatic ductal adenocarcinoma on contrast-enhanced CT images. *Frontiers in Oncology*. 2019; 9:1171
- [11] Huang WC et al. Automatic HCC detection using convolutional network with multi-magnification input images. In: *Proceedings of IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*. Hsinchu, Taiwan: IEEE; 2019. pp. 194-198. DOI: 10.1109/AICAS.2019.8771535
- [12] Dong X, Zhou Y, Wang L, Peng J, Lou Y, Fan Y. Liver cancer detection using hybridized fully convolutional neural network based on deep learning framework. *IEEE Access*. 2020;8: 129889-129898. DOI: 10.1109/ACCESS.2020.3006362
- [13] Zhang Z et al. A novel and efficient tumor detection framework for

- pancreatic cancer via CT images. In: Proceedings of Annual International Conference IEEE Engineering in Medicine and Biology Society. New Orleans, LA, USA: IEEE; 2020. pp. 1160-1164. DOI: 10.1109/EMBC44109.2020.9176172
- [14] Labiadh I, Seddik H, Boubchir L. Deep learning for detection of prostate tumors by microscopic cells and MRI. In: Proceedings of 6th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP). Moncton, Canada and Sfax, Tunisia: IEEE; 2022. pp. 1-5. DOI: 10.1109/ATSIP55956.2022.9805866
- [15] Uysal F, Hardalaç M, Koç M. Classification of T1 and T2 weighted magnetic resonance prostate images using convolutional neural networks. Medical Technologies National Congress (TIPTEKNO). Magusa, Cyprus: IEEE; 2018:1-4. DOI: 10.1109/TIPTEKNO.2018.8596792
- [16] Paul R. Predicting malignant nodules by fusing deep features with classical radiomics features. *Journal of Medical Imaging*. 2018;5(1):011021-1-011021-11
- [17] Huang X, Li Z, Zhang M, Gao S, et al. Fusing hand-crafted and deep-learning features in a convolutional neural network model to identify prostate cancer in pathology images. *Frontiers in Oncology*. 2022;12:1-17
- [18] He S, Ruan J, et al. Combining deep learning with traditional features for classification and segmentation of pathological images of breast cancer. In: 11th International Symposium on Computational Intelligence and Design (ISCID). Hangzhou, China: IEEE; 2018
- [19] Aziz A et al. An ensemble of optimal deep learning features for brain tumor classification. *Computers, Materials & Continua*. 2021;69(2):2653-2670
- [20] Afshar P, Mohammadi A, Konstantinos N. From hand-crafted to deep learning-based cancer radiomics: Challenges and opportunities. arXiv: 1808.07954v3 [cs.CV]. New York, USA: Cornell University; 2019
- [21] Liu S, Xie Y, et al. Pulmonary nodule classification in lung cancer screening with three-dimensional convolutional neural networks. *Journal of Medical Imaging*. 2017;4:4
- [22] Dutta A, Gupta A, Zissermann A: VGG image annotator (VIA). Version: 2.0.9. 2016. Available from: <http://www.robots.ox.ac.uk/vgg/software/via/>
- [23] Deep Learning Toolbox for Matlab. 2022. Available from: <https://it.mathworks.com/help/deeplearning/index.html>
- [24] Torchvision library for Python. 2022. Available from: <https://pytorch.org/vision/stable/index.html>
- [25] Van Der Maaten L, Postma E, Van den Herik L. Dimensionality reduction: A comparative review. *Journal of Machine Learning Research*. 2009;10:66-71
- [26] Kitayama M. Matlab-Kernel-PCA toolbox. 2017. Available from: <https://it.mathworks.com/matlabcentral/fileexchange/7167-matlab-kernel-pca>
- [27] Gaber T, Hassanien T. Particle Swarm Optimization: A Tutorial. Manchester, United Kingdom: IGI Global; 2017
- [28] Too J: Particle Swarm Optimization for Feature Selection. 2022. Available from: <https://github.com/JingweiToo/-Particle-Swarm-Optimization-for-Feature-Selection>



- [29] Waikato Environment for Knowledge Analysis (Weka 3). 2022. Available from: <http://www.cs.waikato.ac.nz/ml/weka/>
- [30] Hall M. Benchmarking attribute selection techniques for discrete class data mining. *IEEE Transactions on Knowledge and Data Engineering*. 2003;15:1-16
- [31] Duda R, Hart R, Stork D. *Pattern Classification*. Hoboken, New Jersey, USA: John Wiley & Sons; 2012
- [32] Meyer-Base A. *Pattern Recognition for Medical Imaging*. Amsterdam, The Netherlands: Elsevier; 2009
- [33] Ojala T, Pietikainen M, Harwood D. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*. 1996; 29(1):51-59
- [34] Chatterjee HS. Various Types of Convolutional Neural Networks. 2019. Available from: <https://towardsdatascience.com/various-types-of-convolutional-neural-network-8b00c9a08a1b>
- [35] Szegedy C, Vanhoucke V, Ioffe S, Shlens S, Wojna Z. Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR6)*. Las Vegas, NV, USA: IEEE; 2016. pp. 2818-2282
- [36] He K et al. Deep residual learning for image recognition. In: *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE; 2016. pp. 770-778
- [37] Huang G, Liu Z, Weinberger KQ. *Densely Connected Convolutional Networks*. CoRR: abs/1608.06993. New York, USA: Cornell University; 2016
- [38] Li Z et al. ConvnextBased fine-grained image classification and bilinear attention mechanism model. *Applied Sciences*. (Basel, Switzerland: MDPI). 2022:1-18