2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Multi-scale patches convolutional neural network predicting the histological grade of hepatocellular carcinoma

Dongsheng Gu, Donghui Guo, Chunwang Yuan, Jingwei Wei, Zhenchang Wang, Hong Zheng and Jie Tian, *Fellow, IEEE*

***Abstract*— Preoperative predicting histological grade of hepatocellular carcinoma (HCC) is a crucial issue for the evaluation of patient prognosis and determining clinical treatment strategies. Previous studies have shown the potential of preoperative medical imaging in HCC grading diagnosis, however, there still remain challenges. In this work, we proposed a multi-scale 2D dense connected convolutional neural network (MS-DenseNet) for the classification of grade. This architecture consisted of three CNN branches to extract features of CT image patches in different scale. Then the outputs for each CNN branch were concatenated to the final fully connected layer. Our network was developed and evaluated on 455 HCC patients from two different centers. For data augmentation, more than 2000 patches for each scale were cropped from transverse section 2D region of interest on these patients. Besides, three-channel inputs including original CT image, tumor region and peritumoral component provided complementary knowledge. Experimental results demonstrated that the proposed method achieved encouraging prediction performance with AUC of 0.798 in testing dataset.**

***Clinical Relevance*—The proposed MS-DenseNet yielded an encouraging prediction performance for HCC histological grade and might assist the clinical diagnosis and decision making of HCC patients.**

I. Introduction

Hepatocellular carcinoma (HCC) is the most common liver cancer with a high incidence of cancer-related death worldwide [1,2]. HCC are always associated with poor prognosis, such as metastasis and early recurrence. Therefore, it is significantly important to acquire early diagnosis and treatment for reducing HCC mortality. The histological grade of HCC has been reported strongly related to survival rate in clinical practice [3]. HCC patients with various degrees of grade can have varying prognosis and treatment strategies. High-grade HCC tumors have a higher risk of recurrence than low-grade tumors. In addition, patients with high histological grade generally need larger safety margins at surgical resection and frequent follow-up after treatment [4,5]. Accordingly, the preoperative accurate assessment of HCC grade is of tremendous value.

Clinically, histological grade of HCC is generally determined by preoperative liver biopsy. However, this process to evaluate HCC grade is not widely accepted for several limitations, including sample error, invasiveness and bleeding [6]. In recent years, noninvasive method extracting high-dimensional features from medical imaging by utilizing machine learning or deep learning algorithm are widely applied in cancer screening, diagnosis and treatment evaluation [7-9]. In terms of grading classification, previous studies confirmed the effectiveness of medical imaging analysis in discriminating high-grade and low-grade in clear renal cell carcinoma, gliomas and pancreatic neuroendocrine [10-12]. Specially, researchers have also explored the potential of medical imaging in predicting HCC grade. Our study is motivated by these successful applications of deep convolution neural network (CNN) in medical imaging analysis [13-15]. Moreover, our proposed multi-scale and multi-channel model increased the accuracy and generalization in predicting histological grade of HCC.

Overall, our contributions are as follows: 1) The proposed multi-scale dense connected convolution neural network (MS-DenseNet) capture the imaging information from region of interest (ROI) in different resolution. This jointly learning process facilitated the classification for histological grade. Simultaneously, the incorporation of multi-scale imaging information can alleviate the "inter-class variations" [16], which exists commonly in clinical medical data. 2) Multi-channel image patches provide the additional prior knowledge of tumor heterogeneity in pixel values (HPV) and heterogeneity in shapes (HS) individually [17]. The three-channel inputs to our model show more effective prediction. 3) This is the first multicenter study to utilize CT image to predict the histological grade of HCC. It demonstrates the effectiveness of our proposed CNN model in different datasets.

This paper is organized as follows. A detailed description of the MS-DenseNet is presented in Section II. Experiment datasets and setup are introduced in Section III. Section IV describes the experimental results. Section V presents the discussion and conclusion.

This paper is supported by the National Natural Science Foundation of China (No. 81227901, 81527805); Ministry of Science and Technology of China (2017YFC1308701, 2017YFC1309100, 2016YFC0102600, 2016YFA0100902, 2016YFC0103803, 2016YFA0201401, 2016YFC0103702, 2014CB748600 and 2016YFC0103001); Chinese Academy of Sciences (No. GJJSTD20170004 and QYZDJ-SSW-JSC005).

[+]D. Gu (gudongsheng2016@ia.ac.cn), J. Wei are with the CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences.

[*]J. Tian (tian@ieee.org) is with the CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University; Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University.

[+]C. Yuan and [*]Z. Wang (cjr.wzhch@vip.163.com) are with the Department of Radiology, Beijing Friendship Hospital, Capital Medical University.

[+]D. Guo and [*] H. Zheng (zhenghong1965@tmu.edu.cn) are with the First Central Clinical College of Tianjin Medical University.

[*] are corresponding authors [+] are co-first authors.

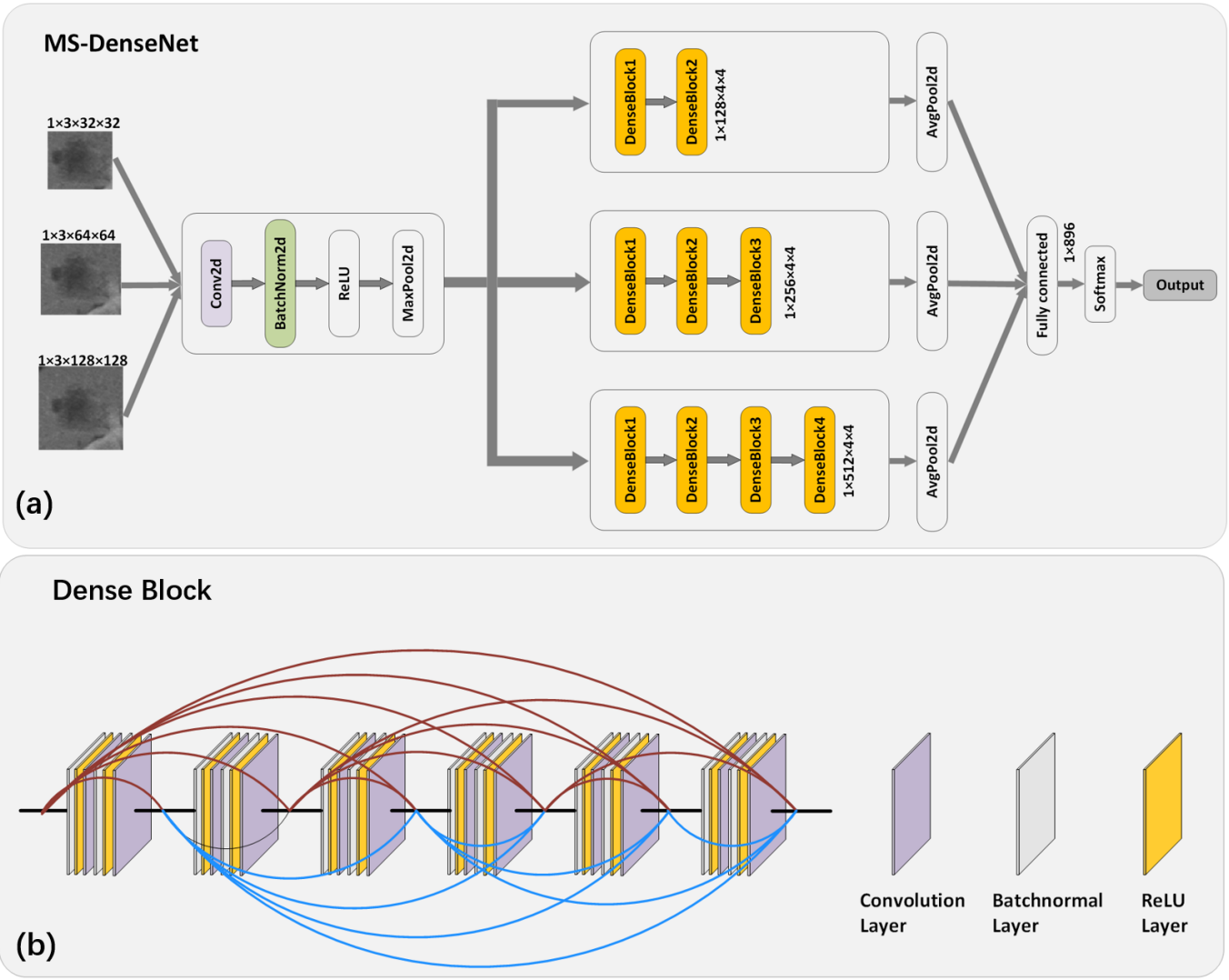

Figure 1. Overview of the proposed method. (a) The architecture for the multi-scale DenseNet. This network contains three branches aiming at capturing features from three different scales and each branch takes a three-channel patch as input. The branches consist of 2, 3 and 4 dense blocks respectively and were finally concatenate to a fully connected layer. (b) The detailed architecture of dense block, including 6 basic bottlenecks which are composed of 2 convolution layers, 2 batch normalization layers and 2 rectified linear unit activate layers.

II. METHOD

A. *MS-DenseNet with multiple channels*

To eliminate the unnecessary noise brought by non-liver tissues and background, we cropped only the tissues within tumor region for imaging analysis. Thus, proper scale for cropping the patches was crucial and influential. Meanwhile, the raw imaging from different scanners and centers could vary in the resolution, which increased the inter-class variations. Therefore, only one-scale raw imaging patches might contribute limited information in the grade classification. It was essential to incorporate information from different scales into our deep neural network. Considering the limited 3D samples, multiple 2D transverse section slice from the internal region of the tumor were collected for modelling. We cropped the 2D patches into three sizes of 32×32, 64×64 and 128×128 pixels manually, which corresponds to the input of three branches of the model.

Our MS-DenseNet is designed based on a classical network architecture DenseNet-121 with pretrained weight on ImageNet [18], which has shown excellent performance in the classification task of natural image. We adopted the major dense-block structure and modified part of layers to adapt our imaging. Figure 1(a) shows the details of our DenseNet. Specifically, the MS-DenseNet incorporated three branches that process patches of different scale respectively. For each branch, we selected the first 2, 3, and 4 dense blocks of the DenseNet-121 with the patch size increasing, which can keep the consistent feature maps after pooling. The 4 dense blocks consisted 6, 12, 24 and 16 basic bottlenecks respectively. The structure of basic bottleneck included two convolution layers of kernel size 1×1 and 3×3, two rectified linear unit (ReLU) activation layers and two batch normalization layers, as is shown in Figure 1(b). The ReLU was used as a nonlinear activation to overcome the vanishing gradient problem and allow models to learn efficiently. Batch normalization layer was applied for reducing internal covariate shift and mitigating overfitting. At the end of CNN model, the three branches were merged through a fully connected layer to calculate the grade probability.

The input patch to each CNN branch was composed of three channels instead of the single channel of raw image. Since the fine tumor boundary was delineated, the prior domain knowledge including HPV and HS can be incorporated to regularize our model [17]. The HPV component of the tumor was generated by setting non-tumor pixels inside the raw image patch to 0, while the HS component was characterized by setting the tumor pixels inside the raw image patch to 0. By setting the pixel of the area of no interest to 0, HPV and HS could focus on the heterogeneity in tumor and shape respectively. Thus, three channels included raw image, HPV and HS component. The knowledge-based three-channel collaborative CNN model characterized the HCC tumor from different aspects and hence yield complementary features.

B. Label smoothing

To further alleviate the overfitting and enhance the generalization ability of the proposed model during training process, we adopted label smoothing technique as regularization [19]. For sample i, (x_i, y_i) had the ground truth distribution p over labels $p(y|x_i)$ and $\sum_{y=0}^{K} p\,(y|x_i) = 1$. K is the candidate label, where $K = 0$ and $K = 1$ represented low-grade and high-grade respectively. Usually the label $p(y|x_i)$ would be a one-hot encoded vector where

$$p(y \mid x_i) = \begin{cases} 1 & if\ y = y_i \\ 0 & otherwise \end{cases}$$

Instead of using one-hot encoded vector, we introduced a prior uniform distribution $u(y|x) = \frac{1}{K}$. The new ground truth label for data (x_i, y_i) would be

$$\begin{aligned} p'(y \mid x_i) &= (1-\varepsilon)p(y \mid x_i) + \varepsilon u(y \mid x_i) \\ &= \begin{cases} 1-\varepsilon+\varepsilon u(y \mid x_i) & if\ y = y_i \\ \varepsilon u(y \mid x_i) & otherwise \end{cases} \end{aligned}$$

Where ε is a weight factor, $\varepsilon \in [0, 1]$, and note that $\sum_{y=0}^{K} p'\,(y|x_i) = 1$. The new ground truth label was employed to replace one-hot label in the cross-entropy loss function as follows:

$$\begin{aligned} L &= -\sum_{i=1}^{n}\sum_{y=0}^{K} p'(y \mid x_i)\log q_\theta(y \mid x_i) \\ &= -\sum_{i=1}^{n}\sum_{y=0}^{K}[(1-\varepsilon)p(y \mid x_i) + \varepsilon u(y \mid x_i)]\log q_\theta(y \mid x_i) \end{aligned}$$

With the regularizer $\varepsilon u(y|x_i)$, we actually prevented the model from predicting too confidently in training dataset. The loss function is minimized during the model training by compute the gradient of L.

III. Experiment

A. Dataset

Our datasets consisted of 455 consecutive patients with pathological confirmed HCC from two hospitals. This dataset was approved by the Institutional Review Board and the requirement for informed consent was waived. CT images of artery phase was acquired from picture archiving and communication systems. HCC grades were retrieved from the pathological report. The manual segmentation for the tumor region was performed by an experienced radiologist slice by slice, using ITK-SNAP software (http://www.itksnap.org/pmwiki/pmwiki.php). Pathological diagnosis results revealed that 287 patients were high-grade corresponding to Edmondson grades II-III, III, III-IV and IV and 168 patients were low-grade corresponding to Edmondson grades I, I-II and II [20]. In this study, we aimed to classify the two-level grade in different datasets.

B. Experimental Setup

To perform a reliable evaluation for the classification experiment, a random partition was conducted on the 455 annotated patients: 273 patients for training, 91 patients for validation and 91 patients for testing at a ratio 3:1:1. We trained the MS-DenseNet on the training set until the loss no longer decreased on the validation set. Finally, the testing set is evaluated using the model. All the training, validation and testing process were performed on 2D patches collected from each patient. The prediction probabilities over the whole patches for one patient were averaged to calculate the probability of patient. Moreover, we augmented the samples by random image translation, rotation and flip operations when training the model, which could help the MS-DenseNet capture invariant tumor features.

During the training process, the weights of the initial network were loaded from pretrained DenseNet-121, and they were updated by the stochastic gradient descent algorithm using a momentum of 0.9 and a batch size of 32. The learning rate was set to 0.1 initially, and decreased by a factor 0.1 every 10 epochs. The ε for label smoothing is set to 0.1. Our MS-DenseNet implementation was based on "pytorch" on a machine with an NVIDIA GeForce GTX 1080Ti GPU.

IV. Result

In order to evaluate the performance of the proposed MS-DenseNet, the area under the curve (AUC), accuracy (ACC), positive predictive value (PPV) and negative predictive value (NPV) were quantitatively calculated.

Quantitative analysis

A single-scale DenseNet with the input of one-channel raw image was chosen as a baseline in this work. A set of ablation experiments were conducted to validate the effectiveness of each component: (a) baseline with input of 32×32 pixels; (b) baseline with input of 64×64 pixels; (c) baseline with input of 128×128 pixels; (d) multi-scale model with the above inputs (MS); (e) (d)+multi-channel input (MS+MC); and (f) (e)+label smoothing, which is also our proposed method (MS-DenseNet). The classification results in testing dataset are shown in Table I.

To investigate the effect of multi-scale information on the classification accuracy, we compared the MS model with single-scale model. As shown in Table 1, the complementary information of the three scales facilitated the classification of HCC grade (AUCs 0.771 vs. 0.736/0.738/0.713). Moreover, the addition of HPV and HS channel proved that the prior domain knowledge could further improve the prediction performance (AUCs 0.783 vs. 0.771). By combing all the

components, the proposed MS-DenseNet performed best in the performance with AUC of 0.798, which was statistically greater than the baseline ($p<0.05$). Meanwhile, both the PPV and NPV were greater than 0.7, confirming the comprehensive ability of the model in stratifying high-grade and low-grade HCC. Receiver operating characteristic curves (ROCs) were plotted for the visualization of different methods (Figure2).

TABLE I.

ABLATION STUDY ON DIFFERENT COMPONENTS

	AUC	*ACC*	*PPV*	*NPV*
Baseline 32×32	0.736	0.692	0.746	0.594
Baseline 64×64	0.738	0.681	0.780	0.561
Baseline 128×128	0.713	0.703	0.630	0.626
MS	0.771	0.714	**0.792**	0.605
MS+MC	0.783	0.725	0.786	0.629
MS-DenseNet	**0.798**	**0.747**	0.758	**0.720**

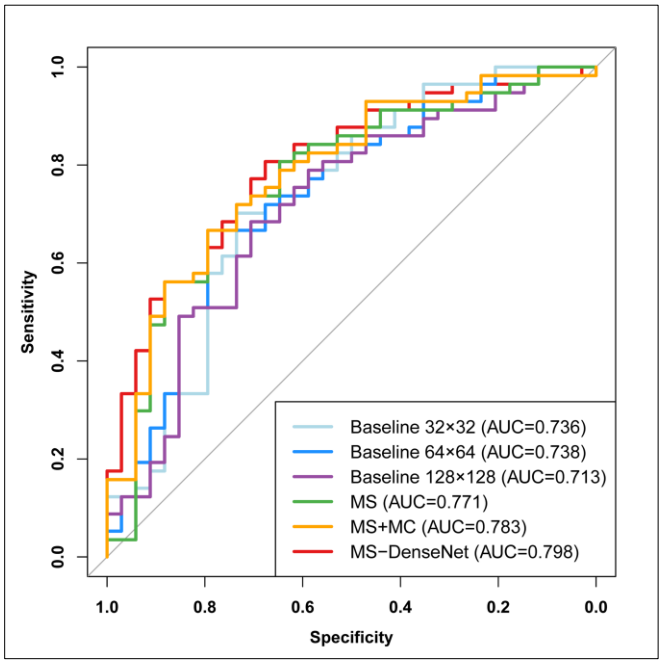

Figure 2. Receiver operating characteristic curves for different methods

V. DISCUSSIONS AND CONCLUSION

In this paper, we proposed a multi-scale 2D dense connected convolutional neural network, incorporating three multi-channel branches to capture the CT imaging information from different scale. In addition, we added label smoothing to alleviate the overfitting and improve the classification performance. Finally, the MS-DenseNet achieved encouraging prediction performance for HCC histological grade over the CT datasets from two centers. In future work, we plan to train the model with multi-phase CT imaging or MR imaging and explore whether they can facilitate the grade classification. Moreover, automatic segmentation algorithm can be used to increase the robustness of our model.

REFERENCES

[1] Forner A, Reig M, Bruix J, "Hepatocellular carcinoma," Lancet, vol. 391, pp. 1301-14, 2010

[2] de Lope CR, Tremosini S, Forner A, Reig M, Bruix J, "Management of HCC," Journal of hepatology, vol. 56 pp. S75-S87, 2012

[3] Nzeako UC, Goodman ZD, Ishak KG. "Comparison of tumor pathology with duration of survival of North American patients with hepatocellular carcinoma," Cancer, vol. 76, pp. 579–588, 1995

[4] Okusaka T, Okada S, Ueno H et al, "Satellite lesions in patients with small hepatocellular carcinoma with reference to clinicopathologic features," Cancer, vol. 95, pp. 1931–1937, 2002

[5] Bruix J, Sherman M, "Management of hepatocellular carcinoma," Hepatology, vol. 42, pp. 1208–1236, 2005

[6] Robert M, Sofair AN, Thomas A, et.al. "A comparison of hepatopathologists' and community pathologists' review of liver biopsy specimens from patients with hepatitis C," Clinical Gastroenterology and Hepatology, vol. 7, pp. 335-338, 2009

[7] Cameron A, Khalvati F, Haider MA, Wong A, "MAPS: a quantitative radiomics approach for prostate cancer detection," IEEE Transactions on Biomedical Engineering, vol. 63, pp. 1145–1156, 2016

[8] Huang YQ, Liang CH, He L et al, "Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer," Journal of Clinical Oncology, vol. 34, pp. 2157–2164, 2016

[9] Liu Z, Zhang XY, Shi YJ et al, "Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer," Clinical Cancer Research, vol. 23, pp. 7253–7262, 2017

[10] Ding JL, Xing ZY, Jiang ZX et al, "CT-based radiomic model predicts high grade of clear cell renal cell carcinoma," European Journal of Radiology, vol. 103, pp. 51–56, 2018

[11] Tian Q, Yan L F, Zhang X, et al, "Radiomics strategy for glioma grading using texture features from multiparametric MRI," Journal of Magnetic Resonance Imaging, vol. 48, pp. 1518-1528, 2018

[12] Gu D, Hu Y, Ding H, et al. "CT radiomics may predict the grade of pancreatic neuroendocrine tumors: a multicenter study," European radiology, pp. 1-11. 2019

[13] Zhou Q, Zhou Z, Chen C, et al. "Grading of hepatocellular carcinoma using 3D SE-DenseNet in dynamic enhanced MR images," Computers in biology and medicine, vol. 107, pp. 47-57, 2019

[14] Zhou W, Wang G, Xie G, et al. "Grading of hepatocellular carcinoma based on diffusion weighted images with multiple b - values using convolutional neural networks," Medical physics, 2019.

[15] Dou T, Zhou W. "2D and 3D convolutional neural network fusion for predicting the histological grade of hepatocellular carcinoma," in the 24th International Conference on Pattern Recognition (ICPR), 2018, pp. 3832-3837

[16] Peng L, Lin L, Hu H, et al. "Multi-scale Residual Network with Two Channels of Raw CT Image and Its Differential Excitation Component for Emphysema Classification," Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, pp. 38-46, 2018

[17] Huang G, Liu Z, Van Der Maaten L, et al. "Densely connected convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp.4700-4708

[18] Xie Y, Xia Y, Zhang J, et al. "Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT,". IEEE transactions on medical imaging, vol. 38, pp. 991-1004, 2018

[19] Szegedy C, Vanhoucke V, Ioffe S, et al. "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp.2818-2826

[20] Edmondson H A, Steiner P E. "Primary carcinoma of the liver. A study of 100 cases among 48,900 necropsies," Cancer, vol. 7, pp. 462-503, 1954