2022 IEEE/ACIS 22nd International Conference on Computer and Information Science (ICIS) | 978-1-6654-9463-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICIS54925.2022.9882470

High precision detection of small hepatocellular carcinoma using improved EfficientNet with Self-Attention

Yuxiang Wang[1]*
School of Electronic and Information Engineering
Tiangong University
Tianjin, China
Euson@outlook.ie

Zhongming Huang[1]*
School of Electronic and Information Engineering
Tiangong University
Tianjin, China
reavenhuang@163.com

***Abstract*—Small hepatocellular carcinoma (SHCC) is among the most fatal cancers, and spotting SHCC symptoms in the early stage is vital for conducting timely treatments. Thus, auxiliary detection algorithms have been developed, especially after convolutional neural networks (CNN) made great progress in processing medical images. However, their performance is confined by dataset, resulting in limitations to accurately detect SHCC appearing small and diffusive in CT images. In our work, Self-Attention mechanism has been introduced as the front end and EfficientNet as our backbone network, contributing to a novel SHCC detection algorithm able to automatically spot subtle lesions through our image-wise annotated dataset in a weakly supervised manner. In our model, the Self-Attention module extracts ROI and background features from original CT images and generates weighed feature map to the EfficientNet. In our backbone network, the EfficientNet learns from input feature maps and is weakly supervised under image-wise annotations. With the pre-process of Self-Attention, our data size for EfficientNet has been reduced, thus enhancing learning efficiency and reducing time consumption. After training on over 1.5k CT images, our model has achieved decent detection performance comparing to other state-of-the-art methods while remaining acceptable complexity.**

Keywords—Small Hepatocellular Carcinoma (SHCC) Detection, Self-Attention, Convolutional Neural Network, Medical Image Processing

I. Introduction

Small hepatocellular carcinoma (SHCC) is accountable for a large number of disease mortalities each year. According to WHO observations [1], SHCC has been the third of major causes of deaths from cancer. Given the perilous risk of exacerbating, it is of high-risk misdiagnosing or ignoring early-stage tumours. To manage to preclude the development of SHCC and conduct therapies earlier, several examine methods were invented, and computed tomography (CT) is widely applied among other techniques. With CT scans, medical workers are able to detect SHCC from observing the darkness of scan images [2]. However, therapists are not always accurate, and early-stage SHCC could be left out [3,4]. Since the early tumours are mostly shaped small and diffusive, they have the potential to escape prudent inspection or cause disagreements between therapists. Fig. 1. shows the SHCC CT images and marks the location of the lesion.

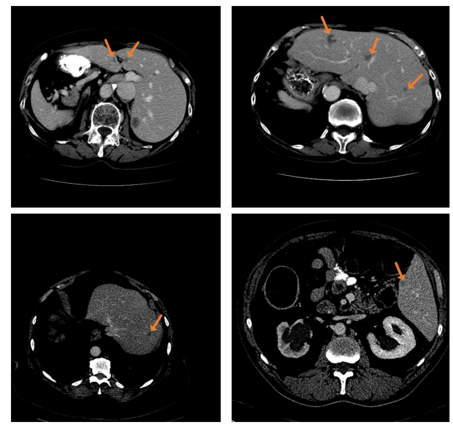

Fig. 1. SHCC CT images and the location of lesions. The lesions of SHCC are usually small which lead to the difficulty of accurate detection.

Predecessors have always been endeavouring to prevent misjudgements of SHCC based on CT scan examinations. At the early stage of computer aided SHCC detection algorithms, simple and mathematical AI approaches were used [5,8], but the overall accuracy was not enough reliable for clinical advising. The advent of convolutional neural network (CNN) further provides researchers a practical and promising solution to process CT scan images [2,3]. By annotating pixel-wise on the CT scan images which have been manually diagnosed, the CNN model is able to learn to distinguish and segment cancerous areas from CT sections. In recent studies of SHCC detection from CT images, Cui et.al. [6] proposed a 3D convolutional neural network; Duc et.al. [7] introduced an adapted convolutional neural network based on U-Net, an originally cellular segmentation-oriented algorithm.

Though prosperous discoveries continue to spring up in SHCC segmentation methods using CNN, knotty problems always occur when those methods are put into practice. To establish a suitable dataset for training traditional CNNs for SHCC detection, huge images should be collected and they are required to be annotated meticulously. This procedure to create a qualified dataset is always exhausting. And because the annotation masks are based on diagnosis from therapists, the judgement of cancerous regions, especially those in early stage, may not be indisputable. Therefore, traditional highly supervised learning is limited by artificial diagnosis, and their

978-1-6654-9463-2/22/$31.00 ©2022 IEEE
ICIS2022, June 26-28, 2022, Zhuhai, China

huge dataset is related to higher computational capacity and longer time consumption.

To further excavate the sensitivity of computer aided algorithms for SHCC CT detection, we would like to explore a novel approach which does not require artificial diagnosis on CT raw data or exhausting pixel-wise annotating. We introduce a novel idea to utilize the attention mechanism and CNN. We adapt the Self-Attention (SA) mechanism [9] into a feature encoder at the front end of CNN backbone, in order to enable our model to automatically learn the region of interest and compare differences between healthy and cancerous liver CT images. For the CNN backbone, we select EfficientNet as the prototype to endow our model high learning efficiency and detection accuracy [10]. Our model is trained on image-wise annotated datasets adapted from the open LiTS benchmark [11], and it gives a classification as the output. The exact detection of SHCC can be decoded from the feature map given by our model. After training, the average detection precision of our model is over 98%, leading congeneric CNNs in SHCC CT segmentation and remaining comparable high accuracy with state-of-the-art methods.

II. Related Work

Computed tomography (CT) examination is an important procedure before the SHCC malady is finally diagnosed [12]. When ultrasonic examination is not enough clear to support analysis of SHCC existence, it becomes necessary to acquire a slice image in the section of abdomen where therapists scrutinize the characteristics inside a liver. Most tumours in a large size are able to be diagnosed by therapists based on CT images. However, in some cases when tumorous shadows in CT scans exist in an ambiguous state, therapists could easily be confused and biopsy may take place. In this case, other risks also lurk [2]. Therapists may ignore early-stage SHCC because of its diffused border or minuscule size, and biopsy is not always practical on patient with fundamental diseases. It is of high risk not able to detect SHCC in time according to the HCC guideline [14].

Therefore, computational approaches aimed to assist SHCC detection based on CT scans were developed. In the earlier stages, machine learning (ML) methods were put into practice [15]. Most ML approaches were based on primary mathematic modelling theories and they do not have fulfilling potentials to handle massive patient data. After the advent of convolutional neural network, the ability for computers to extract patterns in images has been greatly improved. Researchers also adapted achievements in CNN to better suit the specific conditions of SHCC detection. Li W et.al. [16] proposed an CNN automatic SHCC detection based on CT images, their model was trained to segment tumour positions pixel-wise and it is a relatively early implementation in SHCC segmentation algorithms based on a CNN. With delicate labelling of CT image data and sufficient training, such segmentation CNNs have the potential to achieve an accuracy over 70%, and this is not enough for practical diagnosis. As the CNN structure continues to make progress, new models were introduced in SHCC CT detection, especially those of SHCC. Recently, Cui et.al. [6] published a multi-channel 3D-CNN to segment SHCC in CT images and this model was compared between 3D-CNNs of segmentation or non-segmentation sampling, showing robust detection reliability.

However, CNN methods based on traditional supervised learning often demands huge dataset with meticulous labelling. Because the annotations are based on manual diagnosis by therapists, the accuracy and quality of a dataset is confined by the precision of artificial diagnostic advices where mistakes are prone to emerge. Since the attention mechanism came to public, CNN models are endowed with the ability to discriminate subtle differences between sample images in a homogeneous dataset [17]. With the aid of attention mechanism, researchers no longer need to label segmentation datasets in a pixel-wise manner. If the attention mechanism is placed at the front end of a CNN, this weakly supervised learning strategy has the ability to direct the CNN how to segment SHCC regions through comparing mass samples containing SHCC liver CT scans and those which are healthy.

In our model, we utilize the Self-Attention mechanism (SA) as the front end of the CNN backbone. Diao et.al. recently proposed a multi-scale attention mechanism (MSAN-CNN) method in processing whole-slide pathologic images (WPI) [4], contributing to a new method for high resolution medical image processing. But in CT scan examinations, where each image is formatted in grey scale, lower in resolution, and has similar interest regions, we select SA mechanism to achieve lower computational capacity while obtaining decent accuracy. As to the backbone network, we choose to adapt EfficientNet [10] to intake the encoded features from SA front end. After training the backbone on features extracted by SA from SHCC and non-SHCC CT scans, our model turns out to achieve over 98% detection accuracy for SHCC regions, showing improved reliability among traditional SHCC detection methods based on CNN.

III. Our Method

A. Solution

According to the previously mentioned features and problems, we propose to use adapted EfficientNet with Self-Attention [18] to detect SHCC in CT images. Self-Attention mechanism [17] at the front end can effectively screen out the subtle differences in the image and give more "attention" which means our algorithm will focus on analysing these subtle differences and effectively classify them. This Self-Attention module works as a feature extractor and encoder for future learning of the CNN backbone. Meanwhile, the Self-Attention mechanism will automatically select the region of interest, meaning the same parts in a transverse CT section image will be considered as the background while those discriminative parts will be used as the key of our algorithm to learn and analysis. Self-attention mechanism focuses on the blurs and uneven shadows of the liver that may become a tumour, and this mechanism judges irrelevant organs inside and around the liver as the background. Fig. 2. shows where the Self-Attention focused. Through the repeated learning of a large number of SHCC CT images, our algorithm will make more detailed perceptions of the features of SHCC, achieving higher detection accuracy over average therapists according to previous data [19].

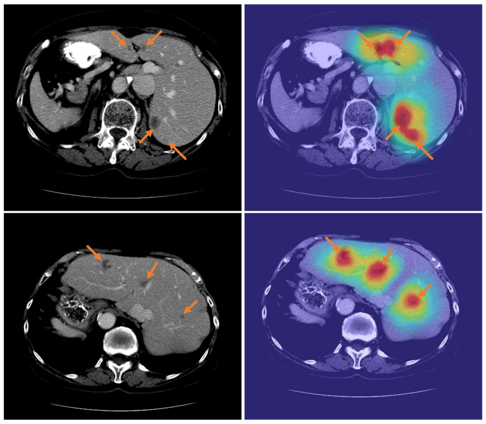

Fig. 2. The locations of SHCC is accurately focused and other locations are marked as the background after the CT image passes through the convolution kernel.

B. Self-Attention Convolution Kernel

Attention mechanism was inspired by human neural science, and self-attention is a derivative of the rudimentary concept of attention [9]. In our implementation of Self-Attention mechanism [20], we assign weights to the pixels of the input images as the pre-processing for latter learning of EfficientNet. We first generate embeddings a of each input image by matrix multiplication, a^i stands for the embedding of the i^{th} input image. Then we extract three attribute vectors Query, Key and Value from the previous embeddings. The three attributes are represented by q^i, k^i, v^i where i is the index of input images in the following formula. Next, we implement self-attention by multiplying the embeddings with Query, Key and Value respectively. Then, by calculating the correlation between Q and K, the matrices Q, K and V are obtained [18]. The weight coefficient of V corresponding to each k is also calculated. The formulas (1)(2)(3) show the above process.

$$q^i = W^q a^i \tag{1}$$

$$k^i = W^k a^i \tag{2}$$

$$v^i = W^v a^i \tag{3}$$

Then, we normalize attention values by referring to SoftMax with the probability distribution with the sum of the weights of all elements as one [18]. According to the probability distribution, the detection characteristics of the attention output matrix can be carried out. The attributes Query, Key and Value are the appropriate parameters learned through separate training, namely weight. By adding these weights to the corresponding pixels which are reduced dimension, the matrix becomes the feature map. The Self-Attention mechanism on our dataset is shown in Fig. 3.

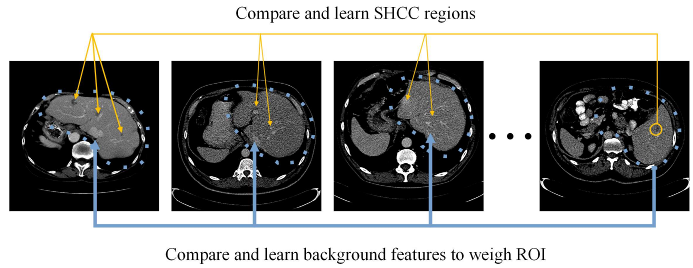

Fig.3. The process of Self-Attention convolution kernel extracting feature. Convolution kernel learns the subject features and background features by comparing a series of pictures.

C. EfficientNet with Self-Attention

The main structure of our algorithm is quite clear. It could be mainly divided into three modules: construct our Self-Attention mechanism, pre-training EfficientNet model [10] and the fusion network. Self-Attention mechanism and EfficientNet model are constructed separately before fusion. After the implementation of these two modules, they will be placed in the fusion connector for encapsulation.

In our whole model where the front mechanism and the backbone CNN are jointed, Self-Attention mechanism functions as an attention convolution kernel. As an instanced kernel stored in the fusion network, its output features can be directly passed to EfficientNet. The existence of this convolution kernel replaces the original feature extraction process originally embedded in EfficientNet. The subtle features of the original data will be extracted one by one and encoded into the "main body", while the same part of each image will be encoded into the "background" through the attention convolution kernel. These two parts of features will be assigned different weights. When the overall weights are summed to one, more weight is assigned to the main body of extracted features while the background part weighs less which could be called heat features map. Through the encoding process by Self-Attention module, we will get a series of image feature vectors that only retain effective detail information and the irrelevant background will not be fed into the EfficientNet. After encoding features, our attention convolution kernel outputs the resulted feature tensor to the pre-training EfficientNet. EfficientNet expands the network depth (Li), network width (Ci) and input resolution (Hi, Wi) on the basis of convolution network [10]. The particle elements of the input heat feature vectors will be magnified by high resolution. And then, they will perform more complex convolution in the extended three-dimensional parameters. EfficientNet will largely improve the training speed and accuracy of our model by sensing and learning based on encoded features. Fig. 4. shows the overall process of data flowing through the improved network.

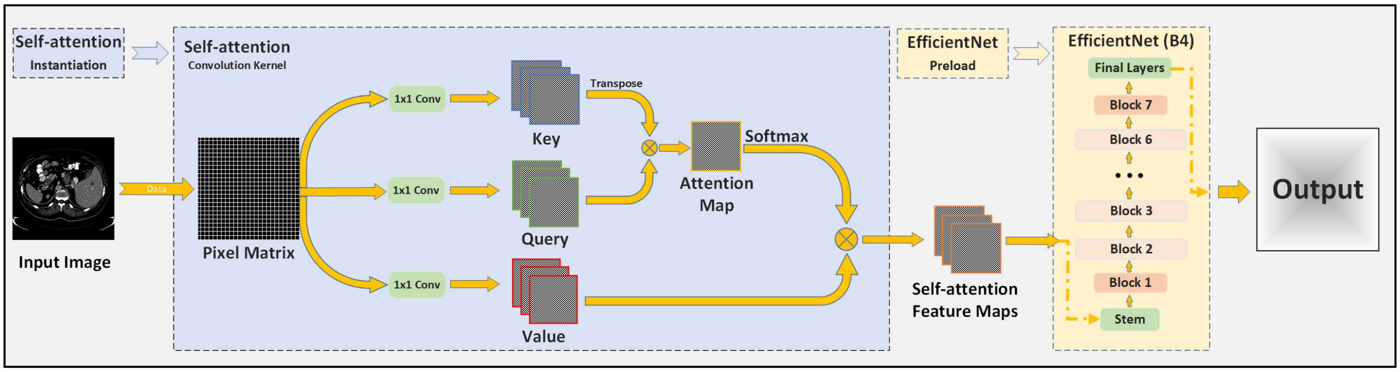

Fig. 4. The whole process of improved network. The original data is extracted from the weighted "main body" and "background" through the Instantiated Self-Attention convolution kernel which will input to EfficientNet for learning in the form of matrix data. Finally, it could get more accurate training and detection results.

IV. Experiment

A. Dataset

Our dataset is selected and re-labelled from the open SHCC CT scan benchmark LiTS [11]. This comprehensive dataset contains three dimensional CT scans of the abdominal region of HCC patients. Totally 130 individuals are covered in LiTS to meet adequate variety in HCC pathology. We selected those CT images taken from the transverse direction, and re-annotated them image-wise into SHCC and non-SHCC groups. Our adapted dataset is consisted of over 10000 image-wise annotated CT scans averagely sampled from each patient's raw data, and we added over 5000 healthy liver section images to provide background contrast for our Self-Attention module.

B. Training and Validation

We divide the dataset into three sets for training, verification and test where their ratio is 3:1:1. The Self-Attention convolution checking accurate feature recognition is a process with continuous comparing and correcting. Therefore, the loss function will fluctuate greatly in the earlier epochs of training. However, as the correct features are focused, loss will be significantly reduced in the following training process. The changes of loss during training are shown as Fig. 5.

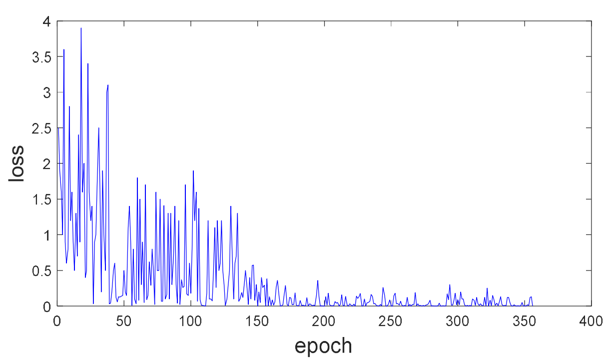

Fig. 5. At the beginning of training, the number of loss is large and fluctuates obviously. After a short period of time, the loss fluctuation becomes smaller. In the middle of the training process, the loss amplitude becomes small and basically remains steady.

Since the accuracy of our detection model is particularly important in providing advice for further detect SHCC. In order to further improve the accuracy of our model, we utilize a small number of the validation set which is about 1500 CT images to evaluate our model and further increase the generalization ability and accuracy of our model. The data of the validation set will also go through all the above processes. After the model is verified, it will be evaluated and corresponding adjustments will be made to further improve the accuracy. The training accuracy changes of mainstream is shown as Fig. 6.

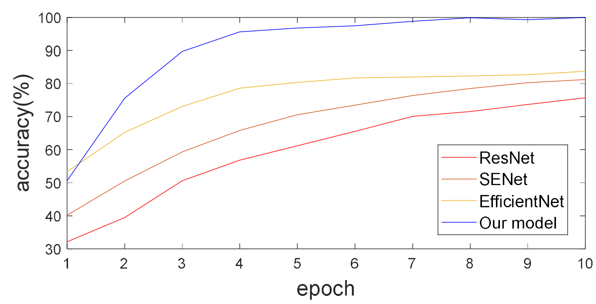

Fig.6. The rising speed of the models using EfficientNet network are significantly higher than the other two networks. And the top accuracy of our model is much higher than other CNNs.

V. Result

A. Top Accuracy

In order to make the accuracy of the algorithm approach near 100%. We have optimized the accuracy of our model in terms of algorithm structure and dataset. Through our experiments, we found that the accuracy of the model can be as high as 98.81%. TABLE I show the accuracy comparison between our model and other mainstream networks on our data set.

TABLE I. Comparison of Top Accuracy of Mainstream CNNs

Detection models	Top Accuracy (%)
ResNet [21]	75.65
SENet [22]	81.24
Original EfficientNet-B4[10]	82.73
Our model	**98.81**

Accuracy is the most important indicator of all auxiliary detection algorithms. In the diagnosis of early SHCC, if we can accurately detect the disease in the early stage of onset, timely therapies would have the opportunity to preclude exacerbation of cancer. Our model has the top accuracy is 98.81% which is close to 100% meaning that our model can detect the occurrence of early cancer with high accuracy, and the fault rate is relatively low. If more data are to be fed in, our accuracy could be further improved. Using this model as the auxiliary method to help detect early liver cancer has high reliability. To a great extent, our method solves the problem that some therapists cannot accurately judge the occurrence of early liver cancer through liver CT examinations. In this respect, it shows that our model has high practical potential in the detection of early SHCC through liver CT images.

B. Parameters and Speed

Usually, parameter complexity and speed are difficult to be equally balanced. However, while adding many parameters to our model, there is no obvious speed decline. The speed of four models running the same dataset is tested under the hardware condition that CPU is i7-10700K and GPU is RTX3060 with CUDA 10.2 and the platform is Windows 10 Pytorch 1.11.0. TABLE II shows the comparison of parameters and speed between our model and other mainstream networks.

Compared with the original EfficientNet, our model has a slight increase in parameters but a small decrease in speed. Nevertheless, our model detection speed remains below 60s which is acceptable in practical application. One reason is the lightweight and high-speed characteristic of EfficientNet, even if we replace the original feature extraction part with a more complex Self-Attention module, with our network parameters to be increased significantly, our whole network can still maintain a fast-running speed. Another reason is that although the Self-Attention module will increase the number of parameters, it will filter out useless information before the pre-trained model starts to learn, meaning the pre-trained model only needs to learn data reduced on dimension to achieve the same result.

TABLE II. COMPARISON OF PARAMETERS AND TEST SPEED

Detection models	Parameters	Speed
ResNet	28.6M	40s
SENet	147.0M	115s
Original EfficientNet-B4	19.2M	32s
Our model	**74.6M**	**49s**

C. Comparison with Recent Researches

We selected four models that were published recently in the field of SHCC CT for comparison. The model based on U-Net proposed by Duc et.al. in 2020 reaches an accuracy of 92.1% [7]. HCCNet was reviewed in 2021, by Wang et.al. which average accuracy can reach 95% [3].The Lim et.al. approach can reach 96% [13].Our model is close in accuracy with those mentioned state-of-the-art methods. In addition, a MSAN-CNN model was introduced to SHCC detection by Diao et.al. in 2022 to solve detection tasks in whole-slide pathologic images (WPI). Their model uses multi-scale attention which achieves an average accuracy of 98.9% in WPI datasets [4]. The average accuracy of our model can be as high as 98.4%, which is almost the same as the MSAN-CNN network using multiscale attention mechanism in lower computational complexities. Obviously, the high accuracy of both attention-based models benefits from the pertinence of attention mechanism for small difference recognition. Although the average accuracy of our model is slightly lower than that of the mentioned MSAN-CNN model, our Self-Attention mechanism has the advantages of easier parallel computing and lower complexity which enable our improved model to have faster training speed and shorter detection period. The TABLE III intuitively shows the comparison of several recent models used for liver cancer detection.

TABLE III. COMPARISON OF AVERAGE ACCURACY OF RECENT MODELS FOR LIVER CANCER DETECTION

Model Name	Average Accuracy (%)	Year
U-Net [7]	92.1	2022
HCCNet [3]	95.0	2021
DCNN in LTP[13]	96.0	2022
MSAN-CNN [4]	**98.9**	**2022**
Our model	**98.4**	**2022**

VI. DISCUSSION

In this research, we used the improved EfficientNet with Self-Attention mechanism to detect small hepatocellular carcinoma (SHCC). From experimental results, it is found that our model has great advantages in dealing with the detection of SHCC. The cooperation of our Self-Attention mechanism and EfficientNet network can accurately identify the lesions of small regions. However, our network also needs to be improved. The main disadvantage is the insufficient stability of our network, which is mainly revealed in the performance fluctuation on different dataset volumes. For CT images with high background similarity in the same group, our network can better pay attention to the location of lesions. Nevertheless, if tested on CT data with varying backgrounds, our network may pay more attention to other irrelevant locations, resulting in a significant reduction in accuracy. These confinements will be improved in our future work.

VII. CONCLUSION

In this paper, we introduce the high-precision detection method of small hepatocellular carcinoma (SHCC) by using improved EfficientNet with Self-Attention model. Through the analysis of pathological data of SHCC, we find that its small area and unclear boundary are the main reasons for the inaccurate detection of SHCC from CT examination. Our model makes targeted optimization for the above challenges. Our Self-Attention mechanism at the front of CNN backbone can directly extract the subtle changes of tumour early occurrence and boundary blurs by comparing the differences of images. Meanwhile, the pre-training network of our improved backbone network has been adapted from

EfficientNet which not only ensures high overall accuracy, but also guarantees the speed of our model. After comprehensive experiments and comparisons, our network accuracy can achieve satisfactory results, which is over 98% much higher than other CNN frameworks by 10% - 30% in the same dataset. In our work, the Self-Attention mechanism is innovatively introduced to SHCC detection, providing a novel effective solution to the problem of accurate detection of SHCC. In recent years, the possibility of SHCC patients being diagnosed or cured has increased continuously. Our work enlightens a new path to solving medical imaging problems.

REFERENCES

[1] Global Cancer Observatory, World Health Organization, https://gco.iarc.fr/. Accessed Apr. 20th , 2022.

[2] Azer SA. Deep learning with convolutional neural networks for identification of liver masses and hepatocellular carcinoma: A systematic review. *World J Gastrointest Oncol. 2019*;11(12):1218-1230. doi:10.4251/wjgo.v11.i12.1218, 2019.

[3] Wang, M., Fu, F., Zheng, B. *et al.* Development of an AI system for accurately diagnose hepatocellular carcinoma from computed tomography imaging data. *Br J Cancer 125* 1111–1121 (2021).

[4] Diao, S., Tian, Y., Hu, W., Hou, J., Lambo, R., Zhang, Z., ... & Qin, W. (2022). Weakly supervised framework for cancer region detection of hepatocellular carcinoma in whole-slide pathologic images based on multiscale sttention convolutional neural network. *The American journal of pathology*, *192* (3), 553-563.

[5] Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. CA Cancer J Clin. 2019;69:127–57.

[6] Cui, H., Wang, K. Y., Li, W. Y., Zhu, H. B., Xiao, L. S., & Liu, L. Artificial Intelligence and Informatics CT images-based 3D convolutional neural network to predict early recurrence of solitary hepatocellular carcinoma after radical hepatectomy, unpublished.

[7] Duc, V. T., Chien, P. C., Chau, T. L. M., Chanh, N. D. T., Soan, D. T. M., Huyen, H. C., ... & Uyen, M. T. T. (2022). Deep learning model with convolutional neural network for detecting and segmenting hepatocellular carcinoma in CT: a preliminary study. *Cureus*, *14* (1).

[8] Jiménez Pérez, M., & Grande, R. G. (2020). Application of artificial intelligence in the diagnosis and treatment of hepatocellular carcinoma: A review. *World journal of gastroenterology*, *26* (37), 5617–5628.

[9] Zhao, H., Jia, J., & Koltun, V. (2020). Exploring self-attention for image recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10076-10085).

[10] Mingxing Tan, Quoc Le, EfficientNet: rethinking model scaling for convolutional neural networks, proceedings of the 36th International Conference on Machine Learning, PMLR 97:6105-6114, 2019.

[11] Patrick Bilic, Patrick Ferdinand Christ, Eugene Vorontsov, Grzegorz Chlebusr et.al., The liver tumor segmentation benchmark (LiTS), arXiv cs.CV 1901.04056, 2019.

[12] Eldad S. Bialecki, Adrian M. Di Bisceglie,Diagnosis of hepatocellular carcinoma, HPB, Volume 7, Issue 1, 2005, Pages 26-34, ISSN 1365-182X.

[13] Sanghyeok Lim, YiRang Shin & Young Han Lee, Arterial enhancing local tumor progression detection on CT images using convolutional neural network after hepatocellular carcinoma ablation: a preliminary study, Scientific Reports (2022) 12:1754 .

[14] Heimbach JK, Kulik LM, Finn RS, Sirlin CB, Abecassis MM, Roberts LR, Zhu AX, Murad MH, Marrero JA. AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatology 2018; 67: 358-380 [PMID: 28130846 DOI: 10.1002/hep.29086].

[15] Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. Gastrointest Endosc 2020 [PMID: 32565184 DOI: 10.1016/j.gie.2020.06.040].

[16] Li W, Jia F, Hu Q. Automatic segmentation of liver tumor in CT Images with deep convolutional neural networks. Computer Communnications 2015; 3: 146-151 [DOI: 10.4236/jcc.2015.311023].

[17] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. Attention is all you need. Advances in neural information processing systems, 30 (2017).

[18] Yuxiang Wang. An improved algorithm of EfficientNet with Self-Attention mechanism. In proceedings of the 2021 International Conference on Internet of Things and Machine Learning, in press.

[19] Bruntha, P.M., Dhanasekar, S., Jency, J.G., Pandian, S.I., Pillai, P., Steven, Pramod, & Malani, V. Performance analysis of certain classifiers for liver CT images, 2019.

[20] Parikh, A. P., Täckström, O., Das, D., & Uszkoreit, J. A decomposable attention model for natural language inference. arXiv preprint arXiv:1606.01933, 2016.

[21] He, K., Zhang, X., Ren, S., & Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778), 2016.

[22] Hu, J., Shen, L., & Sun, G. Squeeze - and - excitation networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.