TwinLiverNet: Predicting TACE Treatment Outcome from CT scans for Hepatocellular Carcinoma using Deep Capsule Networks

C. Pino[1], G. Vecchio[1], M. Fronda[2], M. Calandri[3], M. Aldinucci[4], C. Spampinato[1]

Abstract— **Predicting response to treatment plays a key role to assist radiologists in hepato-cellular carcinoma (HCC) therapy planning. The most widely used treatment for unresectable HCC is the trans-arterial chemoembolization (TACE). A complete radiological response after the first TACE is a reliable predictor of treatment favourable outcome. However, visual inspection of contrast-enhanced CT scans is time-consuming, error prone and too operator-dependent. Thus, in this paper we propose *TwinLiverNet*: a deep neural network that is able to predict TACE treatment outcome through learning visual cue from CT scans. *TwinLiverNet*, specifically, integrates 3D convolutions and capsule networks and is designed to process simultaneously late arterial and delayed phases from contrast-enhanced CTs. Experimental results carried out on a dataset consisting of 126 HCC lesions show that *TwinLiverNet* reaches an average accuracy of 82% in predicting complete response to TACE treatment. Furthermore, combining multiple CT phases (specifically, late arterial and delayed ones) yields a performance increase of over 12 percent points. Finally, the introduction of capsule layers into the model avoids the model to overfit, while enhancing accuracy.**

Clinical relevance— *TwinLiverNet* supports radiologists in visual inspection of CT scans to assess TACE treatment outcome, while reducing inter-operator variability.

Index Terms— **3D Convolutions, Liver Cancer**

I. INTRODUCTION

Hepatocellular Carcinoma (HCC) is the most common liver cancer and the fourth responsible for cancer-related death [1]. HCC early detection can increase the chance of potentially curative treatment that largely depends on tumour characteristics. The most widely used treatment for unresectable HCC is the trans-arterial chemoembolization (TACE) [2], [3], which is recommended for patients with BCLC (Barcelona Clinic Liver Cancer) stage B. Although repeated TACE treatments are often needed, a complete radiologic response after the first TACE session is a predictor of a favorable outcome [4]. However, the radiological examination requires high expertise to inspect visually slice-by-slice CT scans, making the assessment too dependent from the operator. In this paper, we propose an automated tool based on artificial intelligence–techniques for response

prediction to TACE treatment in HCC cases in order to improve treatment allocation. Specifically, the core idea is to build a deep model based on capsule networks [5], [6] for analyzing contrast-enhanced CT scans at multiple stages (i.e., late arterial and delayed phases) and predicting the disease progression under TACE treatments. To our knowledge, automated TACE prediction has not been yet proposed and most of the work on liver cancer has been focused on automated segmentation mainly using either hand-crafted features (e.g., intensity thresholding, region growing, and deformable models) [7], [8], [9], [10] or using deep learned features [11], [12], [13]. The model we propose in this paper - *TwinLiverNet* - is based on 3D convolution kernels to address the existing drawback. Nevertheless, training 3D models is not trivial especially with small datasets, as those in the medical domain, and we tackle the problem of scarce annotated data by extending the representation capability of our model through capsule networks [6], [5]. Furthermore, *TwinLiverNet* is designed to process simultaneously and to learn joint features from multiple CT phases, mainly late arterial and delayed ones.

We test the proposed approach on a dataset consisting of 126 HCC lesions obtaining an average accuracy of 82% in predicting correctly TACE outcome. We also compare *TwinLiverNet* with a single-phase modality (either late arterial or delayed), demonstrating that using both phases at the same time yields to enhanced performance. Ablation studies carried out on the TwinLiverNet, finally, substantiated the made architectural design. In summary, the main contributions of our work are:

- We design and propose the first, to our knowledge, 3D deep model, i.e., for automated TACE response prediction in liver tumor cases using contrast-enhanced CT scans at different phases;
- We explore a multi-phase approach, **TwinLiverNet**, as a variation of the base architecture, to exploit different CT phases peculiarities leading to enhanced accuracy;
- We extend standard 3D convolutional models through capsule networks to effectively deal with 3D data and a limited dataset;

II. RELATED WORK

One of the most investigated tasks in CT data is surely automated segmentation. The standard approach consists of an encoder network, for features embedding, and a decoder network, which produces the segmentation [14]. Although prediction to response treatment is a fairly different task than

*This work was not supported by any organization
[1]PeRCeiVe Lab, Department of Electrical, Electronic and Computer Engineering, University of Catania, Italy `perceive@dieei.unict.it`
[2]Department of Surgical Sciences, Radiology Unit, University of Turin; and Medical Physics Unit, A.O.U. Città della Salute e della Scienza di Torino, Turin, Italy. `marco.fronda87@gmail.com`
[3]Oncology Department, University of Turin, AOU San Luigi Gonzaga, 10043 Orbassano, Italy `marco.calandri@unito.it`
[4]Computer Science Department, University of Torino, Corso Svizzera 185, 10149 Torino, Italy `marco.aldinucci@unito.it`

volume segmentation, encoder networks share many similarities with TwinLiverNet architecture, thus it is the main focus our attention. The CT liver segmentation has been tackled both using handcrafted features and with deep learning based methods. The former category includes thresholding methods [7], [8], deformable models [9], [10], and machine learning based methods [15], [16], [17] relying on traditional algorithms such as support vector machine, random forest and single hidden layer feedforward network. Deep learning methods have gained momentum in CT segmentation, especially since the introduction of U-Net [18]. Indeed, U-NET has become the backbone model for many following architectures, including 3D segmentation with fully convolutional networks (FCNs) [19], [20]. Following this trend, other approaches like Ben-Cohen et al. [21] and Christ et al. [22] proposed FCN-based liver lesion segmentation networks. Sun et al. [23] designed a Multi-Channel Fully Convolutional Network (MC-FCN) to segment liver tumors from multiphase contrast-enhanced CT images. Recently Han [24] proposed a U-Net-like 2.5D FCN which takes a sequence of CT slices and produces the segmentation for the center one. However, the application of deep/machine learning for liver analysis, outside the segmentation field, has been rather limited and mainly using clinical and radiological information for treatment response prediction and survival estimation [25], [26], [27]. One recent approach by Christ et al. [28] has, instead, been proposed for HCC malignancy prediction: the model consists of a lesion segmentation stream through a Cascaded FCN (CFCN), then a 3D neural network, *Survival Net*, predicts the HCC lesion malignancy. To our knowledge, our approach is the first one which tackles the problem of liver tumor response prediction to TACE treatment using only visual cues learned from CT scans. The additional integration of capsule networks [6] in our model is inspired by the very recent trend and related success of these models for medical image understanding [29] also on CT data for cancer node classification [30]. For example, LaLonde et al. [29] proposed a capsule based variant of U-Net for CT scans segmentation, yielding state of the art performance. Recently, Mobiny et al. [30] investigated the use of capsule networks, for automated lung nodule classification from CT data.

III. METHOD

CT data contains both intra-slice and inter-slices information, and approaches that only consider the latter are incomplete as they explicitly ignore dependency between slices. To address this problem, we define the base **LiverNet** architecture: a deep model that consists of a 3D encoder, to learn volumetric information, followed by a 2D encoder to perform (binary) classification (see Fig. 1). Initially, input data (3D crop around a lesion) flows trough a 3D encoder, to exploit inter slice volumetric information. Then, the extracted features are provided to a capsules-based encoder, which predicts treatment outcome as positive or negative. The role of capsule layers is to learn shapes and whole/part relationships from cancer lesions.

Differently from the original CapsNet architecture, *LiverNet* presents an initial 3D encoder of three convolutional layers, to process the input volume. The three 3D convolution layers have respectively 32, 64 and 128, 5x5x5 kernels applied with a stride of 2 in each dimension and followed by a ReLU activation function. The features produced by the 3D encoder is then reshaped, concatenating volumetric features over the channels, and provided as an input to the *primary* capsules. The primary capsules are the lowest layer of the capsules network. This layer has 32 primary capsules, taking as input the features learned through by the 3D convolutional encoder and producing their combinations. The 32 "primary capsules" are similar to convolutional layer in their nature. Each capsule applies 128 9×9 convolutional kernels (with stride 2) to input volume and, therefore, produces $128 \times 3 \times 3$ output tensor. Since there are 32 such capsules, the output volume has a shape of $32 \times 128 \times 3 \times 3$. The output of primary capsules is passed to the output capsules layer, that, given the binary classification nature of our problem, consists of two capsules, one for each class. Each capsule takes as input a $32 \times 128 \times 3 \times 3$ tensor, which can be seen as 1152 input vector. Each of these input vectors has its own weight matrix that maps 128-dimensional input space to the 32-dimensional capsule output space. There are 1152 matrices for each capsule, and also 1152 c_i coefficients and 1152 b_i coefficients used in the dynamic routing. Output capsule layer produces two 32-dimensional vectors v_i, one for each output class. The length of output vector represents the probability of a specific class, being present in the processed data.
The model prediction is defined as

$$pred = \underset{i}{\operatorname{argmax}} \sqrt{\sum_j v_{i|j}^2} \qquad (1)$$

where i is the output capsule index and j is the element index of each element in the ith array. In contrast-enhanced CT, especially for liver cancer analysis, multiple phases of the injection process are identified, namely, non-enhanced, late arterial, portal-venous and delayed phase. Each phase carries out specific information, but the most significant ones, from a radiological perspective, to characterize tumor lesions are the late arterial and delayed phases. In order to exploit this peculiarity, we extend the base LiverNet architecture in order to use both phases for better prediction. We specifically design **TwinLiverNet**(shown in Fig. 2) as a model that jointly learns features from late arterial and delayed visual modalities. **TwinLiverNet** shares the main building blocks with the base LiverNet, but enables to parallel branches, each one processing one specific data modality with the 3D encoder network presented for the one modality counterpart (i.e., LiverNet). The encoded features for each modality are then reshaped and go through a "compression" 2D convolution layer. The compressed features of the two modalities are then concatenated and are fed to a capsule-based classifier, whose architecture is the same as the LiverNet model.
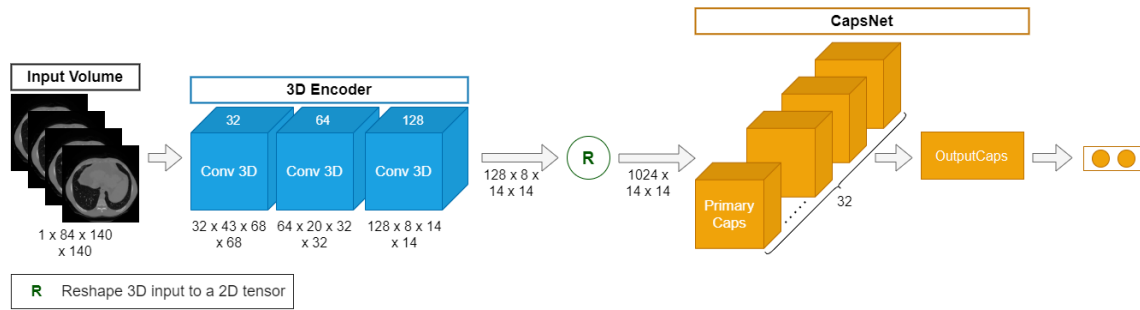
Fig. 1. A simple representation of the proposed architecture. The 3D encoder process the input volume, which is then reshaped and passed to the capsule-network.
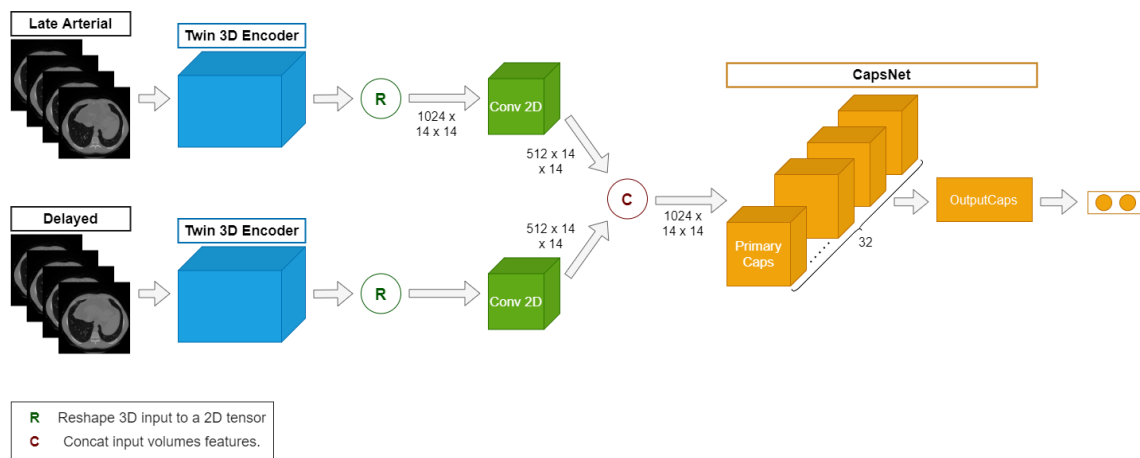

Fig. 2. The architecture for the proposed TwinLiverNet. Two Twin 3D encoders process the input volumes for the different phases. Each features volume is then reshaped and "compressed" by a 2D Conv. The output features are concatenated and passed to the capsule-network.

IV. EXPERIMENTAL RESULTS

A. Dataset

The dataset was provided by the University Radiology Unit of the A.O.U. Città della Salute e della Scienza di Torino (Italy). The dataset consists of 97 CT scans from 92 patients diagnosed with HCC. The number of slices for each scan varies from a minimum of 128 to a maximum of 389. Spatial resolution for each slice is 512×512. Imaging was performed using a CT Optima and a CT Revolution (GE Healthcare). Each contrast-enhanced CT scan consists of four phases: **Non-Enhanced**, **Late Arterial**, **Portal-Venous** and **Delayed**. We consider only late arterial and delayed phases according to radiological criteria for HCC diagnosis and empirical experiments showing limited performance with the Non-Enhanced and Portal-Venous modalities. From each CT scan, multiple lesions are annotated and, in total, our dataset contains 126 HCC lesions that are feed to our LiverNet models for TACE outcome prediction. The inter-slice bounding box size vary from a minimum of 8 to a maximum of 84 slices, while the intra-slice size is between 127 and 139. To uniform data shape, each lesion crop CT was resized to $140 \times 140 \times 90$.

Lesions' annotation is provided in the form of both mRE-CIST (Modified Response Evaluation Criteria in Solid Tumors) classification and lesion bounding box. Annotation was carried out by two experienced radiologists, who labeled each lesion through consensus. mRECIST classes are then grouped into two classes: *positive* in case of complete response (CR) to TACE, and *negative*, in all the other cases (i.e., PR — partial response —, SD — steady disease — and PD — progressive diseases), thus posing the outcome prediction problem as a binary classification task. According to this grouping, the 126 cases are split into 64 positives and 62 negatives. As for evaluation, we employ 5-fold cross validation in order to have, for each fold, 24 lesions (12 positives and 12 negatives). Before lesion cropping, CT scans are pre-processed to enhance visual quality of the images by tweaking brigthness, gamma and contrast.

B. Training procedure

We employ data augmentation mechanisms, namely, random input rotation, translation and mirroring, to reduce overfitting issues. Beside the LiverNet models, we test also

3041

TABLE I

PREDICTION PERFORMANCE OF TESTED MODELS USING 5-FOLD CROSS-VALIDATION. WE REPORT MEAN AND STANDARD DEVIATION OF METRICS COMPUTED OVER THE 5 VALIDATION FOLDS.

Model	Input Phase	Sensitivity	Specificity	Accuracy
Baseline Net (No augm.)	Late Arterial	40.0 ± 6.2	46.7 ± 8.5	43.3 ± 6.7
	Delayed	31.7 ± 6.2	45.0 ± 8.5	38.3 ± 4.0
Baseline Net (Data augm.)	Late Arterial	53.3 ± 8.5	60.0 ± 6.2	56.7 ± 2.0
	Delayed	45.0 ± 4.0	48.3 ± 9.7	46.7 ± 3.1
LiverNet	Late Arterial	71.7 ± 4.0	76.7 ± 8.1	74.1 ± 3.1
	Delayed	63.3 ± 8.5	70.0 ± 8.5	66.7 ± 2.6
TwinLiverNet	Late Arterial + Delayed	$\mathbf{81.7 \pm 8.1}$	$\mathbf{83.3 \pm 7.4}$	$\mathbf{82.5 \pm 7.2}$

a baseline network based on a sequence of 3D and 2D convolutional layers. More specifically, the baseline consists of the LiverNet model without capsule-layers, i.e., the output of the 3D encoder is flattened and fed to two fully-connected layers and a softmax layer for classification. For the baseline we employ binary cross-entropy as loss during training, while for the LiverNet models, we use the length of the instantiation vector for each digit capsule in the last layer and compute the following margin loss, L_k, for each digit capsule, as follows:

$$L_c = T_c \max(0, m^+ - ||v_c||)^2 + \lambda(1 - T_c) \max(0, ||v_c|| - m^-)^2$$
(2)

where $T_c = 1$ if an output of class c is present, $m^+ = 0.9$ and $m^- = 0.1$ and $\lambda = 0.5$. The total loss is thus the sum of all output capsules' losses. All the models are trained from scratch for 200 epochs on a Nvidia Quadro P6000 with 24GB of VRAM. The LiverNet model is trained separately on the *Late Arterial* and the *Delayed* modality, while the TwinLiverNet on both modalities simulataneously. Batch size is set to 24 for the baseline, 24 for LiverNet and 16 for TwinLiverNet. Adam is chosen as optimizer algorithm using a learning rate of $5 * 10^{-6}$ and $[0.9, 0.999]$ as β parameters. As metrics to test the goodness of the devised models for TACE outcome prediction, we measure sensitivity, specificity and accuracy.

C. Results

We conduct few experiments to substantiate our architectural design. In particular, we perform control study to verify the effectiveness of a) data augmentation, b) capsule layers and c) using multiple phases images. Thus, we test our baseline with and without data augmentation, the LiverNet model with data augmentation when using only either late arterial images or delayed images, and the model (Twin-LiverNet) when using simultaneously both modalities. The achieved results (mean and standard deviation over the 5 folds) are reported in Table I and show that: 1) using data augmentation yields to significantly improved results; 2) the late arterial phase alone is more informative than the delayed

one and, as such, it allows for better prediction, 3) capsule layers are necessary to correctly predict treatment outcome with an accuracy gain of about 20 percent points w.r.t. the baseline, and 4) when using both modalities together in the TwinLiverNet model, the performance gains is about of 10 percent points yielding a final accuracy in the prediction of 82.5% (with a sensitivity of 81.7% and a specificity of 83.3%).

V. CONCLUSIONS

Correctly estimating a response to cancer treatment, in particular hepatocellular carcinoma, plays a central role in therapy planning, supporting medical staff in the decision making process. This work proposes the first, to our knowledge, deep learning solutions that attempt to predict the outcome to TACE treatment by using only visual information learned from pre-treatment CT scans. We initially introduce a traditional 3D CNN architecture, and then enhance its representational capabilities through capsule networks. Capsules are used to enforce spatial morphology learning for prediction, while coping with limited high-dimensional training data. The resulting model, LiverNet, is then tested when using a) single phase data at a time, and b) late arterial and delayed phase images (TwinLiverNet). Performance analysis, carried out on a dataset of 126 HCC lesions, shows that capsule layers are necessary to predict correctly TACE outcome yielding a 82.5% accuracy, significantly higher than purely convolution-based architectures. The highest performance gain is obtained when using both late arterial and delayed images demonstrating that, correlations between the two data modalities, allows for a better characterization of HCC tumor lesions. Furthermore, among the different phases of contrast-enhanced CT scanning, the late arterial is the most informative one. In conclusion, these results obtained in this preliminary work open interesting perspectives on visual characterization of cancer lesions through deep learning. Such a characterization may go beyond the only HCC here presented, but it can applied to a variety of other cancers, such as genomics profiling of lung cancers [31].

REFERENCES

[1] Ju Dong Yang, Pierre Hainaut, Gregory J. Gores, Amina Amadou, Amelie Plymoth, and Lewis R. Roberts. A global view of hepato-cellular carcinoma: trends, risk, prevention and management. *Nature Reviews Gastroenterology & Hepatology*, 16(10):589–604, Oct 2019.

[2] J-L Raoul, B Sangro, A Forner, V Mazzaferro, F Piscaglia, L Bolondi, and Riccardo Lencioni. Evolving strategies for the management of intermediate-stage hepatocellular carcinoma: available evidence and expert opinion on the use of transarterial chemoembolization. *Cancer treatment reviews*, 37(3):212–220, 2011.

[3] Joong-Won Park, Minshan Chen, Massimo Colombo, Lewis R Roberts, Myron Schwartz, Pei-Jer Chen, Masatoshi Kudo, Philip Johnson, Samuel Wagner, Lucinda S Orsini, et al. Global patterns of hepatocellular carcinoma management from diagnosis to death: the bridge study. *Liver International*, 35(9):2155–2166, 2015.

[4] Beom Kyung Kim, Seung Up Kim, Kyung Ah Kim, Yong Eun Chung, Myeong-Jin Kim, Mi-Suk Park, Jun Yong Park, Sang Hoon Ahn, Man Deuk Kim, Sung Il Park, et al. Complete response at first chemoembolization is still the most robust predictor for favorable outcome in hepatocellular carcinoma. *Journal of hepatology*, 62(6):1304–1310, 2015.

[5] Geoffrey E Hinton, Alex Krizhevsky, and Sida D Wang. Transforming auto-encoders. In *International conference on artificial neural networks*, pages 44–51. Springer, 2011.

[6] Sara Sabour, Nicholas Frosst, and Geoffrey E Hinton. Dynamic routing between capsules. In *Advances in neural information processing systems*, pages 3856–3866, 2017.

[7] Jan Hendrik Moltz, Lars Bornemann, Volker Dicken, and H Peitgen. Segmentation of liver metastases in ct scans by adaptive thresholding and morphological processing. In *MICCAI workshop*, volume 41, page 195, 2008.

[8] Luc Soler, Hervé Delingette, Grégoire Malandain, Johan Montagnat, Nicholas Ayache, Christophe Koehl, Olivier Dourthe, Benoit Malassagne, Michelle Smith, Didier Mutter, et al. Fully automatic anatomical, pathological, and functional segmentation from ct scans for hepatic surgery. *Computer Aided Surgery*, 6(3):131–142, 2001.

[9] Damon Wong, Jiang Liu, Yin Fengshou, Qi Tian, Wei Xiong, Jiayin Zhou, Yingyi Qi, Thazin Han, S Venkatesh, and Shih-chang Wang. A semi-automated method for liver tumor segmentation based on 2d region growing with knowledge-based constraints. In *MICCAI workshop*, volume 41, page 159, 2008.

[10] Daniel Jimenez-Carretero, Laura Fernandez-de Manuel, Javier Pascau, Jose M Tellado, Enrique Ramon, Manuel Desco, Andres Santos, and Maria J Ledesma-Carbayo. Optimal multiresolution 3d level-set method for liver segmentation incorporating local curvature constraints. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3419–3422. IEEE, 2011.

[11] Kumardeep Chaudhary, Olivier B Poirion, Liangqun Lu, and Lana X Garmire. Deep learning–based multi-omics integration robustly predicts survival in liver cancer. *Clinical Cancer Research*, 24(6):1248–1259, 2018.

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

[13] Amirhossein Kiani, Bora Uyumazturk, Pranav Rajpurkar, Alex Wang, Rebecca Gao, Erik Jones, Yifan Yu, Curtis P Langlotz, Robyn L Ball, Thomas J Montine, et al. Impact of a deep learning assistant on the histopathologic classification of liver cancer. *NPJ digital medicine*, 3(1):1–8, 2020.

[14] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.

[15] Weimin Huang, Yongzhong Yang, Zhiping Lin, Guang-Bin Huang, Jiayin Zhou, Yuping Duan, and Wei Xiong. Random feature subspace ensemble based extreme learning machine for liver tumor detection and segmentation. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4675–4678. IEEE, 2014.

[16] Eugene Vorontsov, Nadine Abi-Jaoudeh, and Samuel Kadoury. Metastatic liver tumor segmentation using texture-based omni-directional deformable surface models. In *International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging*, pages 74–83. Springer, 2014.

[17] Trong-Ngoc Le, Hieu Trung Huynh, et al. Liver tumor segmentation from mr images using 3d fast marching algorithm and single hidden layer feedforward neural network. *BioMed research international*, 2016, 2016.

[18] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[19] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.

[20] Xiaomeng Li, Hao Chen, Xiaojuan Qi, Qi Dou, Chi-Wing Fu, and Pheng-Ann Heng. H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes. *IEEE transactions on medical imaging*, 37(12):2663–2674, 2018.

[21] Avi Ben-Cohen, Idit Diamant, Eyal Klang, Michal Amitai, and Hayit Greenspan. Fully convolutional network for liver segmentation and lesions detection. In *Deep learning and data labeling for medical applications*, pages 77–85. Springer, 2016.

[22] Patrick Ferdinand Christ, Mohamed Ezzeldin A Elshaer, Florian Ettlinger, Sunil Tatavarty, Marc Bickel, Patrick Bilic, Markus Rempfler, Marco Armbruster, Felix Hofmann, Melvin D'Anastasi, et al. Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3d conditional random fields. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 415–423. Springer, 2016.

[23] Changjian Sun, Shuxu Guo, Huimao Zhang, Jing Li, Meimei Chen, Shuzhi Ma, Lanyi Jin, Xiaoming Liu, Xueyan Li, and Xiaohua Qian. Automatic segmentation of liver tumors from multiphase contrast-enhanced ct images based on fcns. *Artificial intelligence in medicine*, 83:58–66, 2017.

[24] Xiao Han. Automatic liver lesion segmentation using a deep convolutional neural network method. *arXiv preprint arXiv:1704.07239*, 2017.

[25] Hyun Jeong Park, Jung Hoon Kim, Seo-youn Choi, Eun Sun Lee, Sang Joon Park, Jae Young Byun, and Byung Ihn Choi. Prediction of therapeutic response of hepatocellular carcinoma to transcatheter arterial chemoembolization based on pretherapeutic dynamic ct and textural findings. *American Journal of Roentgenology*, 209(4):W211–W220, 2017.

[26] Wen-Hsien Ho, King-Teh Lee, Hong-Yaw Chen, Te-Wei Ho, and Herng-Chia Chiu. Disease-free survival after hepatic resection in hepatocellular carcinoma patients: a prediction approach using artificial neural network. *PLoS one*, 7(1):e29179, 2012.

[27] Jie Peng, Shuai Kang, Zhengyuan Ning, Hangxia Deng, Jingxian Shen, Yikai Xu, Jing Zhang, Wei Zhao, Xinling Li, Wuxing Gong, et al. Residual convolutional neural network for predicting response of transarterial chemoembolization in hepatocellular carcinoma from ct imaging. *European radiology*, 30(1):413–424, 2020.

[28] Patrick Ferdinand Christ, Florian Ettlinger, Georgios Kaissis, Sebastian Schlecht, Freba Ahmaddy, Felix Grün, Alexander Valentinitsch, Seyed-Ahmad Ahmadi, Rickmer Braren, and Bjoern Menze. Survivalnet: Predicting patient survival from diffusion weighted magnetic resonance images using cascaded fully convolutional and 3d convolutional neural networks. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 839–843. IEEE, 2017.

[29] Rodney LaLonde, Ziyue Xu, Sanjay Jain, and Ulas Bagci. Capsules for biomedical image segmentation. *arXiv preprint arXiv:2004.04736*, 2020.

[30] Aryan Mobiny and Hien Van Nguyen. Fast capsnet for lung cancer screening. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 741–749. Springer, 2018.

[31] Francesca Trenta Francesca Cordero Ulas Bagci Franceco Rundo Sebastiano Battiato Daniela Giordano Marco Aldinucci Concetto Spampinato Carmelo Pino, Simone Palazzo. Interpretable deep model for predicting gene-addicted non-small-cell lung cancer in ct scans. In *International Symposium on Biomedical Imaging (ISBI 2021)*, 2021.