Journal of Cancer Research and Clinical Oncology (2025) 151:84
https://doi.org/10.1007/s00432-025-06119-8

RESEARCH

An interpretable ensemble model combining handcrafted radiomics and deep learning for predicting the overall survival of hepatocellular carcinoma patients after stereotactic body radiation therapy

Yi Chen[1,2] · David Pasquier[1,3,4] · Damon Verstappen[1] · Henry C. Woodruff[1,5] · Philippe Lambin[1,5]

Received: 24 June 2024 / Accepted: 23 January 2025 / Published online: 14 February 2025

Abstract

Purpose Hepatocellular carcinoma (HCC) remains a global health concern, marked by increasing incidence rates and poor outcomes. This study seeks to develop a robust predictive model by integrating radiomics and deep learning features with clinical data to predict 2-year survival in HCC patients treated with stereotactic body radiation therapy (SBRT).

Methods This study analyzed a cohort of 186 HCC patients who underwent SBRT. Radiomics features were extracted from CT scans, complemented by collection of clinical data. Training and validation of machine learning models were conducted using nested cross-validation techniques. Deep learning models, leveraging various convolutional neural networks (CNNs), were employed to effectively integrate both image and clinical data. Post-hoc explainability techniques were applied to elucidate the contribution of imaging data to predictive outcomes.

Results Handcrafted radiomics features demonstrated moderate predictive performance, with area under the receiver operating characteristic curve (AUC) values ranging from 0.59 to 0.72. Deep learning models, harnessing the fusion of image and clinical data, exhibited improved predictive accuracy, with AUC values ranging from 0.71 to 0.81. Notably, the ensemble model, amalgamating handcrafted radiomics and deep learning features with clinical data, demonstrated the most robust predictive capability, achieving an AUC of 0.86 (95% CI: 0.80–0.93).

Conclusion The ensemble model represents a significant advancement, providing a comprehensive tool for predicting survival outcomes in HCC patients undergoing SBRT. The inclusion of interpretability methods such as Grad-CAM enhances transparency and understanding of these complex predictive models.

Keywords Hepatocellular carcinoma · Stereotactic body radiation therapy · Radiomics · Handcrafted features · Deep learning

Yi Chen and David Pasquier contributed equally to this work.

✉ David Pasquier
d-pasquier@o-lambret.fr

1 The D-Lab, Department of Precision Medicine, GROW—Research Institute for Oncology and Reproduction, Maastricht University, Maastricht, The Netherlands

2 Engineering Research Center of Text Computing & Cognitive Intelligence, Key Laboratory of Intelligent Medical Image Analysis and Precise Diagnosis of Guizhou Province, State Key Laboratory of Public Big Data, College of Computer Science and Technology, Ministry of Education, Guizhou University, Guiyang 550025, People's Republic of China

3 Academic Department of Radiation Oncology, Centre O Lambret, Lille, France

4 University of Lille, Centrale Lille, CNRS, UMR 9189-CRIStAL, Lille, France

5 Department of Radiology and Nuclear Imaging, GROW - Research Institute for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands

Introduction

Hepatocellular carcinoma (HCC) is the seventh most common cancer and the third most common cause of cancer-related deaths worldwide. The incidence of HCC has been rising worldwide over the last 20 years and is expected to increase until 2030 in some countries, including the United States (IARC 2022, Kulik et al. 2019). Infection by the hepatitis B or C virus are the main risk factors for HCC development, although non-alcoholic steato-hepatitis associated with metabolic syndrome or diabetes mellitus is becoming a more frequent risk factor in the West (Llovet et al. 2021). The pathophysiology of HCC is complex, with the interplay of factors including genetic predisposition, interactions between viral and non-viral risk factors, the cellular microenvironment, and the severity of the chronic liver disease (Llovet et al. 2021). Treatment options depend on tumor burden, severity of liver dysfunction, and performance status. The prognosis is severe, and the Barcelona Clinic for Liver Cancer (BCLC) staging system, which pairs these parameters with a recommended therapy, is the most widely used (Vogel et al. 2021). Less than 30% of the patients are eligible to surgery, mainly because of tumor burden and/or liver dysfunction. Trials and series over the past two decades have shown the efficacy and safety of radiotherapy and stereotactic body radiation therapy (SBRT) in HCC, leading to its adoption in some recent guidelines (Vogel et al. 2021, Lewis et al. 2022).

Radiological data from routine imaging procedures could potentially be used to non-invasively assess tumor characteristics. To step beyond subjective interpretation, the field of radiomics has emerged, characterized by the high-throughput extraction of quantitative features from medical imaging, which can then be correlated with biological and clinical endpoints (Lambin et al. 2012; Gillies et al. 2016). Radiomics allows for the discovery of non-invasive imaging biomarkers that could improve HCC diagnosis and prediction of response to treatment. Handcrafted radiomics (HCR) used pre-defined features in this process, and most published work in this area focuses on diagnosis, prediction of vascular invasion after surgery, and prognosis after local treatment such as surgery and trans arterial chemo embolization (TACE). Kim et al. demonstrated that a combination of clinical (Child-Pugh score, serum alpha fetoprotein, tumour size) and radiomics features (surface area-to-volume ratio, kurtosis, median, and size zone variability) can predict post-TACE overall survival; the combined model was a better predictor of survival ($p<0.0001$) than the clinical score model or the radiomics score model only (Kim et al. 2018). Similar findings were also reported in other studies, in which radiomics features extracted from pre-treatment imaging (CT or MRI) for the prediction of treatment response after TACE were compared to post-treatment response evaluation (Yuan et al. 2019).

Data concerning the prediction of survival after intensity-modulated radiotherapy or SBRT are scarce (Wu et al. 2020; Harding et al. 2021, Wei et al. 2021; Maino et al. 2024). On a different topic, Shen et al. used clinical, radiomics, and dose-volumetric factors to predict the occurrence of SBRT-related radiation-induced liver disease in HCC patients (Shen et al. 2022).

Deep learning (DL), a sub-field of radiomics that dispenses the need for handcrafted features at the expense of needing more data for training, has also shown potential in predicting diagnosis and outcomes after surgery for HCC. In a retrospective series, DL models achieved accurate pre-operative prediction of vascular invasion and might facilitate predicting tumor recurrence and mortality after surgery (Yang et al. 2022). DL and HCR s may have a complementary value, potentially establishing a more robust tool (Parekh et al. 2019, Beuque et al. 2023). Therefore, this study aimed to apply machine learning methods that combine HCR and deep learning image features models from CT scans, and clinical features to predict a 2-year survival probability for patients treated with SBRT.

While currently radiomics algorithms are still considered to be “black boxes”, because of the perceived lack of interpretability, Gradient-weighted class activation maps (Grad-CAM) are post-hoc interpretability methods that are useful for understanding the results of DL models, (Lundberg et al. 2017, Selvaraju et al. 2017; Salahuddin et al. 2022). Grad-CAM uses the gradient from the last convolution layer to produce a coarse localized map for highlighting important regions of the image for prediction or segmentation.

Materials and methods

Patient database

A total of 318 consecutive and inoperable HCC patients were recruited from June 2007 to December 2018, who underwent SBRT procedures using Cyberknife® at the O. Lambret Center. The diagnosis of HCC was made using liver biopsy or imaging, according to Barcelona diagnosis criteria (Vogel et al. 2021). Indication of SBRT was based on the decision by the institutional multidisciplinary consultation committee of hepatic tumors, which considered among other factors, patient comorbidities, ECOG score, anatomical relations of the HCC, and multiplicity and site of lesions. Follow-up was planned every three months in the first three years, and every six months afterwards. OS was determined from the start of treatment until death from any cause. Full inclusion details and clinical results of this

series are described in Roquette et al. (2022). The study complies with the "reference methodology" adopted by the French Data Protection Authority (CNIL), all patients provided informed consent for data collection and analysis and the study was approved by the relevant Institutional Ethics Committee. The association between the risk of death and prognostic factors was analyzed using a Cox model after checking the proportional hazard assumption. The software used for the statistical analysis was R version 4.1.2 (R Core team 2021).

Computed tomography

Previous work from our team indicated a notable impact of CT phase on HCC radiomics feature values (Ibrahim et al. 2021a). Consequently, we included 186 out of 318 patients from whom a CT scan in the arterial contrast phase was available. The late arterial phase with bolus tracking was used. The same CT-scanner was used in all the patients (Toshiba-MEC CT3). Matrix and slice thickness were 512 × 512 and 1 mm in all the patients respectively. Mean (SD) pixel spacing was 1.07 (±0.18) mm. The tube voltage was 120 kV in 77% of the patients and 135 kV in 23% of the patients. Mean (SD) tube current by slice was 329 (±105) mA. The FC13 kernel was used for all but 4 patients, for whom the FC03 kernel was used (same kernel with beam hardening correction) (Kitagawa et al. 2010). Patients' characteristics are presented in the supplementary data (Table S1).

The contours of the HCC lesions were manually delineated by a radiation oncologist with more than 20 years of experience. Patients underwent MRI, which was registered with CT to help with delineation of the grosstumor volume (GTV). All the GTV contours were checked by a second radiation oncologist with more than 15 years of experience. Clinical Target Volume (CTV5) was defined as GTV + 5 mm isotropic margins, bounded by the liver contour. A second CTV (10 mm margin, CTV10) was created using the software package MIM v7.1.3 (MIM Software, Cleveland, OH). Healthy liver volume was defined as liver minus CTV10. Peri-tumoral volume (ring) was defined as the volume of liver contained within the 5 mm expansion around the tumor. The largest lesion was studied in patients with multiple intrahepatic tumors.

The liver was delineated according to Radiation Therapy Oncology Group guidelines, excluding large vessels and the gallbladder (Jabbour et al. 2014). The automatic liver contours have been generated by the software package MIM v7.1.3 (MIM Software, Cleveland, OH), checked and modified if necessary by a radiation oncologist with more than 15 years of experience.

Handcrafted radiomics model preprocessing and training

All images underwent resampling to a uniform voxel size of 1 × 1 × 1 mm³, rectifying variations in pixel dimensions, given that each scan possessed a consistent slice thickness of 1 mm. Subsequent to this standardization, distinct data preprocessing methodologies were employed. In the context of handcrafted radiomics, discretization of image intensity was performed utilizing a fixed bin width of 25 Hounsfield Units (HU). For deep learning applications, image normalization was achieved through Z-score standardization, followed by cropping of images according to liver contours. Feature extraction was performed according to the Image Biomarker Standardization Initiative (IBSI) guidelines (Zwanenburg et al. 2020). A total of 100 IBSI features were extracted from each segmentation. The features can be divided into first-order intensity, histogram statistics, shape, and texture features. A full list and a description of the features can be found in the PyRadiomics documentation (Radiomics features 2019).

For the HCR and clinical feature models (HCR&CL_ML), we applied a nested cross-validation approach with 10 inner and 10 outer folds. The outer loop accounts for bias in the train-test splits and the inner loop is used for hyperparameter tuning by applying a grid search. The feature selection steps are laid out in the appendix. If a feature gets selected more than five times in the training folds of the outer loop, it is included in the model. The hyperparameters with the best ROC AUC over all inner loop folds are selected for the outer loop (Figure S1). To avoid "data leakage" by including "test" data during feature selection, features are selected for each of the outer training sets.

Deep learning model preprocessing and training

For the deep learning model, we first selected the best deep learning model based on the image only; we tried to train different variants of ResNet, DesNet, Mobilenet, and EffecientNet models. The ResNet50 showed the best performance and was set as the baseline model for all deep learning models with images. Then we applied the same ten-fold cross-validation selection like the Handcrafted model part, and the patient selection was the same as the HCR pipeline. To evaluate which factors will affect the prediction result, we create a model based on the image only, on a combination of images and clinical features, and on the image plus handcraft radiomics and clinical features, respectively. DenseNet, EfficientNet, and ResNet were employed as foundational 3D classification architectures each comprising distinct sub-models tailored to the model's depth. This selection also encompassed a comprehensive survey

of convolutional neural network (CNN) models within the Monai framework. Notably, all classification networks shared a common fully connected layer for conventional categorization tasks, serving as a classifier. However, given our specific binary classification objective, we necessitated the replacement of the original fully connected layer with a binary classification equivalent.

We selected clinical features through multifactorial Cox analysis to construct two deep learning models: one combining clinical features with images, and another combining clinical features, radiomics, and images. Following this, preprocessed images were input into the ResNet50 model, while clinical data and HCRs were concurrently processed using a linear model. These outputs were then integrated as novel inputs for a secondary linear model. The computational training of these models was executed on three GPUs, specifically NVIDIA GeForce RTX 4090 Ti, each equipped with 24 GB of RAM. During the training phase, the Adam optimizer was employed, with an initial learning rate set at 0.0001 and a weight decay of 0.01 to facilitate L1 normalization. Additionally, the learning rate was subjected to a StepLR decay schedule.

The ensemble of deep learning and radiomics was obtained by the mean value of the predicted probabilities of the HRD&CL_ML model and DL model (Fig. 1).

Results

Clinical features

The median survival of the 186 patients was 21.5 months (95%CI:19–28). Most of the clinical features that we evaluated were correlated; bilirubin, albumin levels and ascites are taken into account in the Child-Pugh score, Child-Pugh score and presence of portal thrombosis are included in the BCLC-score. Consequently, including all the factors potentially associated with poor prognosis in a multivariate model would not be reliable even though these factors are not overlapping totally. As in the article describing clinical outcomes in the general population (Roquette et al. 2022), we evaluated the prognostic value of each factor identified as significant in univariate analysis when adjusted for Child Pugh and BCLC score. In this multivariate analysis, Albi score and bilirubin are no longer significantly associated with poor survival when adjusted for Child Pugh score ($p=0.6$ and $p=0.3$, respectively). Ascitis remains associated with prognosis when adjusted for Child Pugh score ($p=0.014$). Child Pugh score, GTV and the GTV/ Healthy Liver ratio remained significantly associated with prognosis when adjusted for BCLC too ($p<0.05$). The BCLC score remained significantly associated with poor survival when adjusted for each of these factors separately. Results of

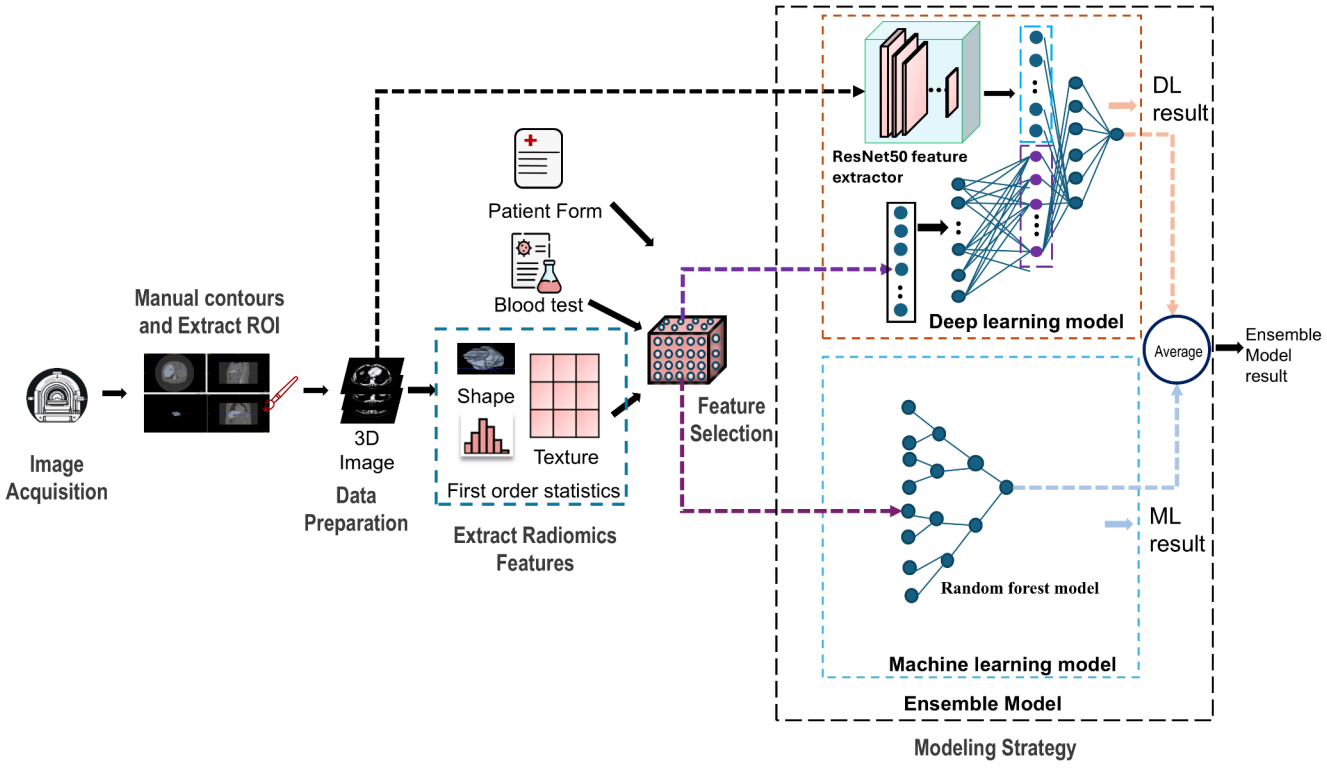

Fig. 1 Representation of the workflow of constructing models

univariate and multivariate analysis are presented in Supplementary data (Table S2).

Handcrafted radiomics

In order to conduct training in the same way, we train deep learning and machine learning models based on 10-fold cross-validation, using the same folds for both models. Differentiation in handcrafted radiomics features (HCR) values is evident across various Regions of Interest (ROI) including the GTV, CTV, and healthy liver tissue. To highlight these distinctions, we extracted features from each ROI. Our analysis reveals that these ROI-based features do not significantly enhance the predictive performance compared to the Barcelona Clinic Liver Cancer (BCLC) staging system, which remains the standard predictive model.

Upon further examination, the largest GTV volume consistently demonstrated the most stable HCRs across all ROIs. Clinical features showed a better result than all results based on the image radiomics features. Finally, we integrated the largest tumor volume with the most effective clinical feature-based model yielding a slight improvement in the mean Receiver Operating Characteristic (ROC) Area Under the Curve (AUC) and a reduction in the standard deviation (SD) during nested cross-validation (nCV) (Fig. 2 and Figure S2). A SHapley Additive exPlanations (SHAP) values for the clinical and gross tumor volume model is presented in Figure S3.

Deep learning model

In the context of a deep learning framework, a series of ablation studies elucidates the relationship among various inputs, all of which are based on the same initial conditions, including training hyperparameters and image preprocessing. The results of different network architectures are presented in Table 1. Firstly, we train the ResNet50 model based images. Then, we add clinical and HCRs by step based on the ResNet50. Finally, the model ensembling clinical and radiomics features demonstrated superior performance in 10-fold cross-validation compared to the model relying solely on image inputs and to the model with images inputs, clinical and handcraft radiomics. However, when both image and clinical features were integrated and processed using

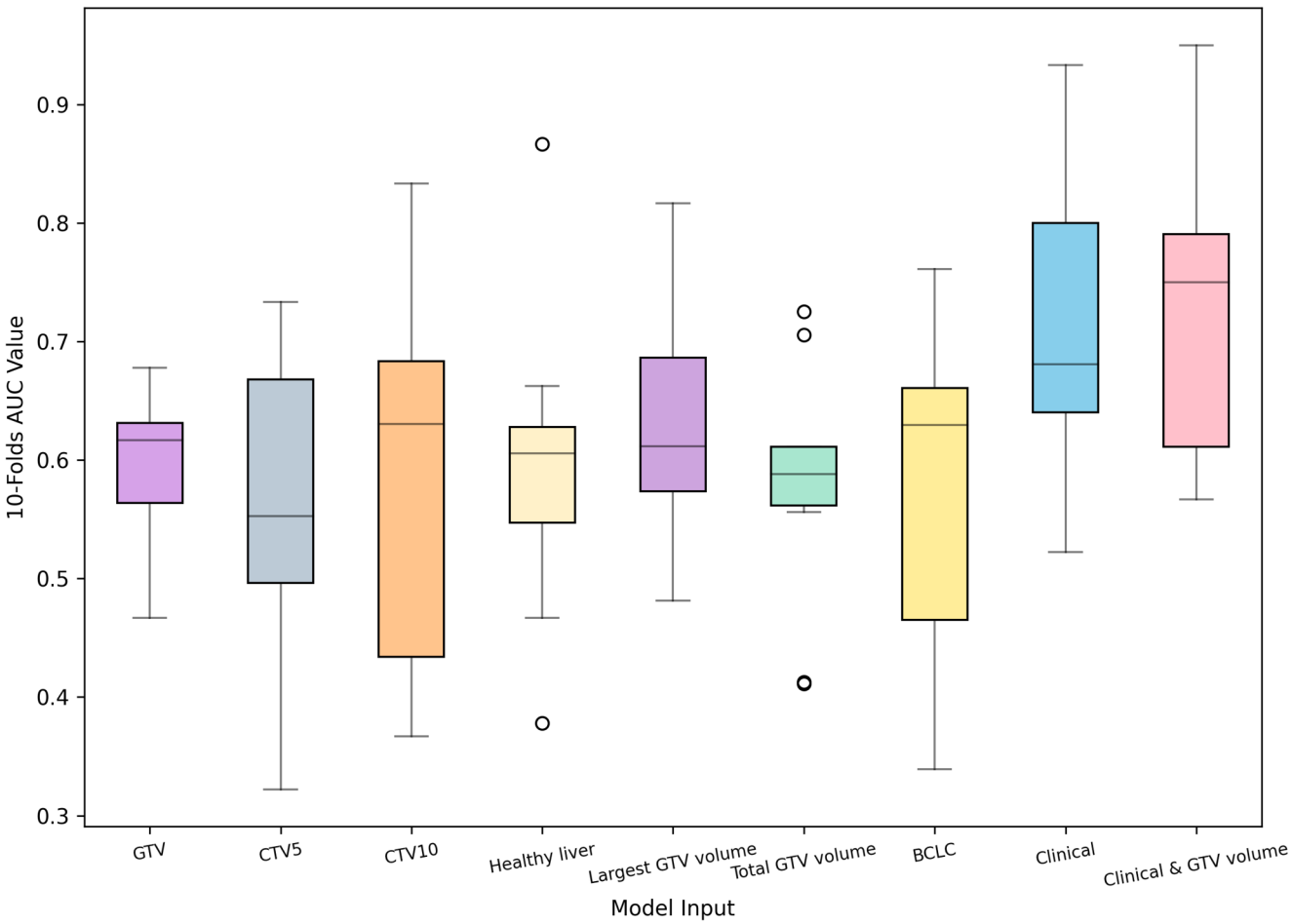

Fig. 2 Boxplot of the ROC AUC for handcrafted radiomics and clinical features

Table 1 Deep learning results in different architectures based on the images

Architectures	AUC
ResNet50	0.59 (CI95: 0.59–0.82)
ResNet101	0.51 (CI95:0.49–0.53)
DensNet121	0.49 (CI95:0.48–0.52)
EfficientNetB0	0.48 (CI95: 0.45–0.52)
EfficientNetB7	0.50 (CI95:0.48–0.54)

a modified ResNet50-based model, the results were nearly analogous to those of the solely clinical linear model. Notably, the model that combined image, clinical, and radiomics features exhibited the most commendable performance, as evidenced by its superior mean ROC AUC when compared with the models mentioned above (Table 2). It also shows the radiomics features can implement the deep features we automatically extract from the deep learning model.

Ensemble model

In the ensemble model, we obtain the probability by calculating the average probability from the machine learning and deep learning models in the same patient. The AUC of the ensemble model is 0.86 (95CI:0.8–0.93) (Fig. 3).

Interpretability

Gradient-Class Activation Maps (Grad-CAM) explanation was applied to the last convolutional layer of the DLmodel for each test case by highlighting the part of the input image that the model considers important for prediction in the form of a heat map (Selvaraju et al. 2017). Grad-CAM can be used for interpretability to visualize which part of the image the model is focusing on. In Fig. 4, we applied the Grad-CAM to show which part will contribute to the classification result. The red and green parts indicate negative and positive contributions to the final classification result respectively. The activated areas are in the tumor and the healthy liver.

Table 2 Deep learning results

Model Input	Clinical	Image	Image+clinical	Image+radiomics	Image+clinical+radiomics
Average AUC	0.71	0.59	0.76	0.53	0.81
SD	0.12	0.08	0.08	0.07	0.13
Range	0.55–0.84	0.5–0.78	0.59–0.82	0.4-0.633	0.63–0.94
CI95	[0.61–0.76]	[0.5–0.67]	[0.69–0.82]	[0.48–0.58]	[0.8–0.93]

AUC Area under the curve; SD: Standard deviation; CI95: Confidence Interval 95%

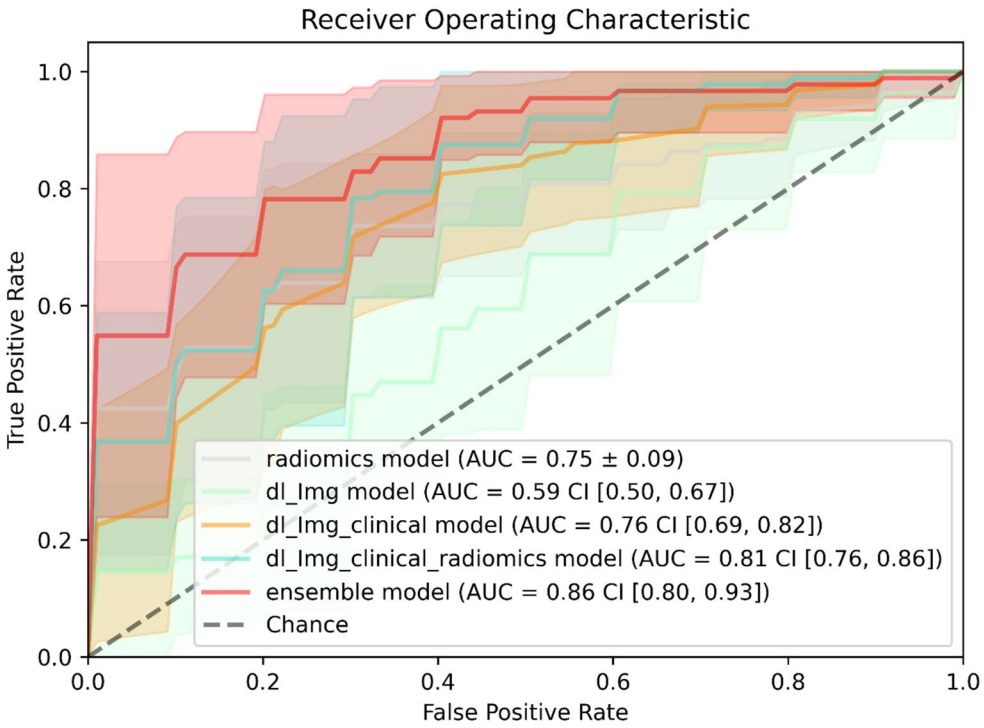

Fig. 3 ROC AUC curves for deep learning (DL) based on radiomics, images (Img) +/- clinical features, and ensemble models

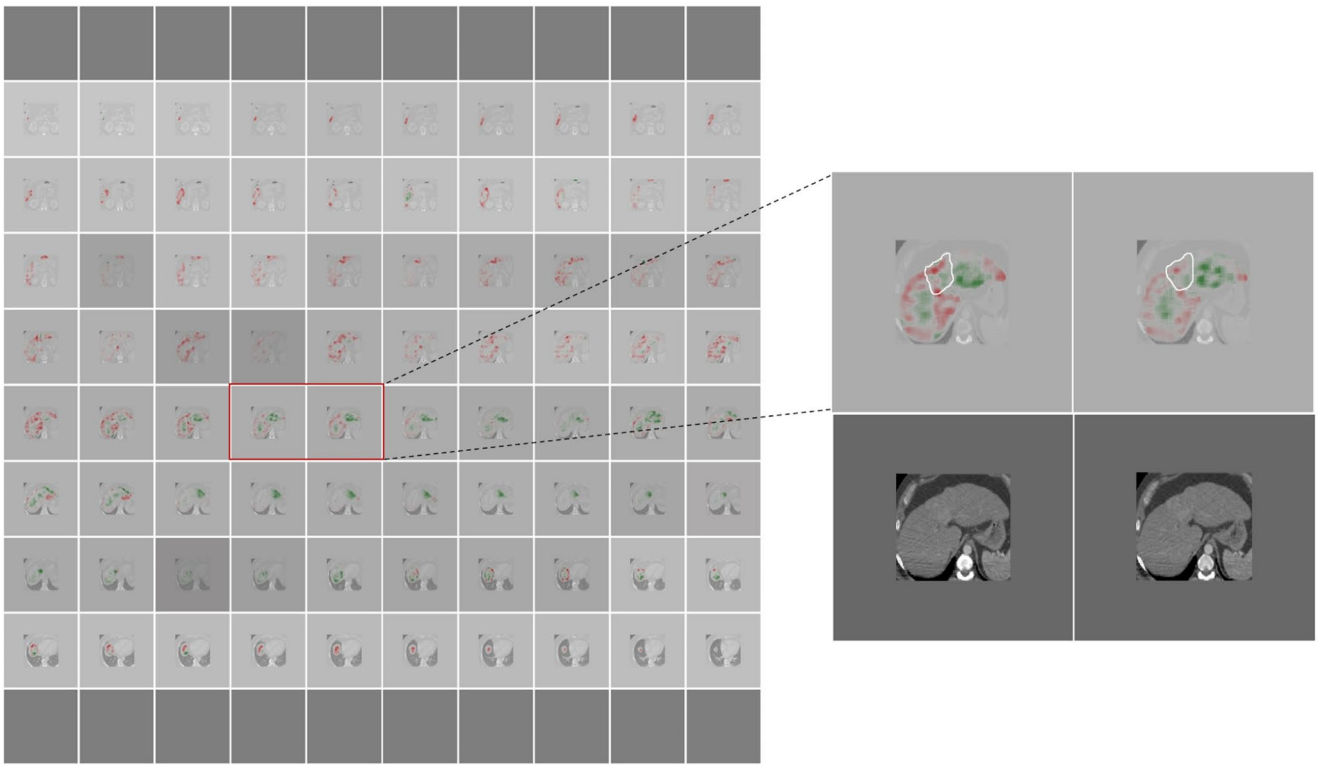

Fig. 4 Gradient-Class Activation Maps in the whole liver of a patient using deep learning model (red and green: negative and positive contribution respectively). White contour: gross tumour volume

Discussion

Data concerning the prediction of survival after radiotherapy in HCC patients are scarce. In this study, we proposed an ensemble model combining HCR and DL based on images and clinical features that can be useful for predicting survival in HCC patients treated with SBRT. This Ensemble model allows for a high AUC of 0.86 (95CI:0.8–0.93), higher than all the other models, including the BCLC, which is the clinical reference (Vogel et al. 2021).

Our HCR model has shown the importance of tumor volume, which is considered a radiomics feature, insofar as this variable is not used in clinical practice. Cozzi et al. appraised the ability of a radiomics-based analysis to predict local response and overall survival in 106 patients with HCC treated with intensity-modulated radiotherapy, without validation cohort. A radiomics signature, made by one single feature was finally identified: after elastic net regularization, the most significant covariates were compacity and BCLC stage, with only compacity significant to Cox model fitting (Cozzi et al. 2017).

One of the strengths of our study lies in the use of CT scans performed in the same injection phase. Our team showed that CT phase has a significant impact on HCC radiomics feature values. Ibrahim et al. showed that the majority of HCR were not reproducible between the arterial and portal venous phases before or after ComBat harmonization (Ibrahim et al. 2021a). The injection phase is not specified in some series (Wu et al. 2020; Kim et al. 2018; Yuan et al. 2019; Harding et al. 2021, Maino et al. 2024), while working on the arterial and venous phases could alter reproducibility.

We worked on images from the same scanner in all patients, avoiding reproducibility issues. Indeed generalizability of radiomics signatures is affected significantly by differences in scan acquisition and reconstruction settings. In contrast to what has been previously reported, the previous results of our team demonstrated that ComBat cannot be applied to all HCR but rather on a percentage of those -the "ComBatable" HCR- which differed depending on the data being harmonized (Ibrahim et al. 2021b). However working on images from a single scanner may hinder generalization of the results.

The vast majority of the series are from Asia, regions where the etiology of HCC is largely dominated by viral hepatitis (Harding et al. 2021, Wu et al. 2020; Yuan et al. 2019; Shan et al. 2019). Our database includes patients

whose etiology is variable, due to lifestyle habits (alcoholic cirrhosis), viral etiology, hoping for better generalization.

All causes of death were studied in our study, without accounting for competing risks related to deaths not associated with tumor progression. Indeed, the cause of death can sometimes be difficult to determine in these patients, as liver failure, for example, may occur with or without a link to tumor progression. Additionally, we studied features of the 'healthy liver' for a holistic approach, making the method relevant.

The Grad-CAM allowed us to visualize the active areas that led to the DL result. These areas are not only tumoral, in connection with the pathophysiology and the prognosis of HCC. The prognosis is severe, driven by cirrhosis, metastasis, and subsequent HCC. The healthy (non-tumoral) liver is the site of fibrosis, the site of production of immunomodulatory cytokines promoting local recurrence and the occurrence of other HCC lesions, as well as liver failure (Llovet et al. 2021; Affo et al. 2017). Wei et al. used features from both gross tumor volume (GTV) regions and liver exclusive of the GTVs in 167 patients (Wei et al. 2021). Radiomics, clinical and raw image deep learning network (DNN) models were combined to predict the risk probability for OS. The final models yielded c-indices of 0.650 (95%CI: 0.635–0.683) for combined models in a nested cross-validation scheme (Wei et al. 2021). This method, which also shows the role of the non-tumour liver, is close to ours, although our AUC results appear to be higher. Finally, the combination of HCR and DL allows us to increase the predictive value of radiomics, as demonstrated in previous articles concerning breast mammograms or diagnosis of Idiopathic Pulmonary Fibrosis (Beuque et al. 2023; Refaee et al. 2022). The limitations of our study are its single-center nature, without external cohort, so multicentric external validation is needed to verify its generalizability before taking steps towards clinical application.

Interpretability methods like GradCAM can sometimes produce visually compelling explanations that may not accurately reflect the underlying model or data-generating process. These methods might lack sensitivity to important aspects of the model, and their effectiveness can be undermined by invariances, such as model or label randomization (Adebayo 2018).

In conclusion, the ensemble model combining HCR and DL based on images and clinical features can be useful for predicting 2-year survival in patients treated with SBRT offering improved model interpretability through SHAP and Grad-CAM. It may serve as a useful method to help post-operative follow-up and management. Further prospective multi-center studies should be considered for validation.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00432-025-06119-8.

Author contributions Conception and design: DP, HW, PL; provision of study materials or patients: DP; collection, data analysis and interpretation: YC, DP, DV; HW and PL guided the data analysis and processing; manuscript writing: YC, DP, DV; final approval of manuscript: all authors.

Funding This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability The data that support the findings of this study are not openly available and are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate This research study was conducted retrospectively from data obtained for clinical purposes. The study protocols were in compliance with the French Regulation for retrospective studies (MR004).

Consent for publication Not applicable.

Consent to participate Consent for participation was obtained in compliance with the French Regulation for retrospective studies.

Competing interests The authors declare no competing interests.

References

Adebayo J, Gilmer J, Muelly M, Goodfellow I, Hardt M, Kim B (2018) Sanity Checks for Saliency Maps. 32nd Conference on Neural Information Processing Systems (NeurIPS), Montréal, Canada https://arxiv.org/pdf/1810.03292

Affo S, Yu LX, Schwabe RF (2017) The role of Cancer-Associated fibroblasts and fibrosis in Liver Cancer. Annu Rev Pathol 12:153–186

Beuque MPL, Lobbes MBI, van Wijk Y, Widaatalla Y, Primakov S, Majer M, Balleyguier C, Woodruff HC, Lambin P (2023) Combining Deep Learning and Handcrafted Radiomics for classification of suspicious lesions on contrast-enhanced mammograms. Radiology 307:e221843

Cozzi L, Dinapoli N, Fogliata A, Hsu WC, Reggiori G, Lobefalo F, Kirienko M, Sollini M, Franceschini D, Comito T, Franzese C, Scorsetti M, Wang PM (2017) Radiomics based analysis to

predict local control and survival in hepatocellular carcinoma patients treated with volumetric modulated arc therapy. BMC Cancer 17:829

Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278:563–577

Harding-Theobald E, Louissaint J, Maraj B, Cuaresma E, Townsend W, Mendiratta-Lala M, Singal AG, Su GL, Lok AS, Parikh ND (2021) Systematic review: radiomics for the diagnosis and prognosis of hepatocellular carcinoma. Aliment Pharmacol Ther 54:890–901

Ibrahim A, Widaatalla Y, Refaee T, Primakov S, Miclea RL, Öcal O, Fabritius MP, Ingrisch M, Ricke J, Hustinx R, Mottaghy FM, Woodruff HC, Seidensticker M, Lambin P (2021a) Reproducibility of CT-Based Hepatocellular Carcinoma Radiomic features across different contrast imaging phases: a proof of Concept on SORAMIC Trial Data. Cancers 13:4638

Ibrahim A, Refaee T, Leijenaar RTH, rimakov S, Hustinx R, Mottaghy FM, Woodruff HC, Maidment ADA, Lambin P (2021b) The application of a workflow integrating the variable reproducibility and harmonizability of radiomic features on a phantom dataset. PLoS ONE 16:e0251147

International Agency of Research on Cancer Cancer today [Internet]. [acceded 2024 May 09th]. Available from: http://gco.iarc.fr/today/home

Jabbour SK, Hashem SA, Bosch W, Kim TK, Finkelstein SE, Anderson BM, Ben-Josef E, Crane CH, Goodman KA, Haddock MG, Herman JM, Hong TS, Kachnic LA, Mamon HJ, Pantarotto JR, Dawson LA (2014) Upper abdominal normal organ contouring guidelines and atlas: a Radiation Therapy Oncology Group consensus. Pract Radiat Oncol 4:82–89

Kim J, Choi SJ, Lee SH, Lee HY, Park H (2018) Predicting survival using pretreatment CT for patients with hepatocellular carcinoma treated with transarterial chemoembolization: comparison of models using radiomics. AJR Am J Roentgenol 211:1026–1034

Kitagawa K, George RT, Arbab-Zadeh A, Lima JAC, Lardo AC (2010) Characterization and correction of beam-hardening artifacts during dynamic volume CT assessment of myocardial perfusion. Radiology 256:111–118

Kulik L, El-Serag HB (2019) Epidemiology and management of Hepatocellular Carcinoma. Gastroenterology 156:477–491e1

Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJWL (2012) Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 48:441–444

Lewis S, Dawson L, Barry A, Stanescu T, Mohamad I, Hosni A (2022) Stereotactic body radiation therapy for hepatocellular carcinoma: from infancy to ongoing maturity. JHEP Rep InnovHepatol 4:100498

Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, Lencioni R, Koike K, Zucman-Rossi J, Finn RS (2021) Hepatocellular carcinoma. Nat Rev Dis Primer 7:1–28

Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, CA: Curran Associates Inc. pp. 4768–77

Maino C, Vernuccio F, Cannella R, Franco PN, Giannini V, Dezio M, Pisani AR, Blandino AA, Faletti R, De Bernardi E, Ippolito D, Gatti M, Inchingolo R (2024) Radiomics and liver: where we are and where we are headed? Eur J Radiol 171:111297

Parekh VS, Jacobs MA (2019) Deep learning and Radiomics in Precision Medicine. Expert Rev Precis Med Drug Dev 4:59–72

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ Acceded February 04th

Radiomic, Features (2019) — Pyradiomics V3.0.1.Post9+Gdfe2c14 Documentation Available at:https://pyradiomics.readthedocs.io/en/latest/features.html. Acceded February 04th

Refaee T, Salahuddin Z, Frix AN, Yan C, Wu G, Woodruff HC, Gietema H, Meunier P, Louis R, Guiot J, Lambin P (2022) Diagnosis of idiopathic pulmonary fibrosis in high-resolution computed Tomography scans using a combination of handcrafted Radiomics and Deep Learning. Front Med 9:915243

Roquette I, Bogart E, Lacornerie T, Ningarhari M, Bibault JE, Le Deley MC, Lartigau EF, Pasquier D, Mirabel X (2022) Stereotactic body Radiation Therapy for the management of Hepatocellular Carcinoma: efficacy and safety. Cancers (Basel) 14:3892

Salahuddin Z, Woodruff HC, Chatterjee A, Lambin P (2022) Transparency of deep neural networks for medical image analysis: a review of interpretability methods. ComputBiol Med 140:105111

Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-cam: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE International Conference on Computer Vision. Cambridge, MA: IEEE. pp. 618–26

Shan QY, Hu HT, Feng ST, Peng ZP, Chen SL, Zhou Q, Li X, Xie XY, Lu MD, Wang W, Kuang M (2019) CT-based peritumoral radiomics signatures to predict early recurrence in hepatocellular carcinoma after curative tumor resection or ablation. Cancer Imaging 19:11

Shen PC, Huang WY, Dai YH, Lo CH, Yang JF, Su YF, Wang YF, Lu CF, Lin CS (2022) Radiomics-based predictive model of Radiation-Induced Liver Disease in Hepatocellular Carcinoma patients receiving stereo-tactic body Radiotherapy. Biomedicines 10:597

Vogel A, Martinelli E, (2021) Updated treatment recommendations for hepatocellular carcinoma (HCC) from the ESMO Clinical Practice Guidelines Ann Oncol 32(6):801–805. https://doi.org/10.1016/j.annonc.2021.02.014

Wei L, Owen D, Rosen B, Guo X, Cuneo K, Lawrence TS, Ten Haken R, El Naqa I (2021) A deep survival interpretable radiomics model of hepatocellular carcinoma patients. Phys Med 82:295–305

Wu K, Shui Y, Sun W, Lin S, Pang H (2020) Utility of Radiomics for Predicting Patient Survival in Hepatocellular Carcinoma with Portal Vein Tumor thrombosis treated with stereotactic body Radiotherapy. Front Oncol 10:569435

Yang Y, Zhou Y, Zhou C, Ma X (2022) Deep learning radiomics based on contrast enhanced computed tomography predicts microvascular invasion and survival outcome in early stage hepatocellular carcinoma. Eur J Surg Oncol 48:1068–1077

Yuan C, Wang Z, Gu D, Tian J, Zhao P, Wei J, Yang X, Hao X, Dong D, He N, Sun Y, Gao W, Feng J (2019) Prediction early recurrence of hepatocellular carcinoma eligible for curative ablation using a Radiomics nomogram. Cancer Imaging 19:21

Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R, Bogowicz M, Boldrini L, Buvat I, Cook GJR, Davatzikos C, Depeursinge A, Desseroit MC, Dinapoli N, Dinh CV, Echegaray S, El Naqa I, Fedorov AY, Gatta R, Gillies RJ, Goh V, Götz M, Guckenberger M, Ha SM, Hatt M, Isensee F, Lambin P, Leger S, Leijenaar RTH, Lenkowicz J, Lippert F, Losnegård A, Maier-Hein KH, Morin O, Müller H, Napel S, Nioche C, Orlhac F, Pati S, Pfaehler EAG, Rahmim A, Rao AUK, Scherer J, Siddique MM, Sijtsema NM, Socarras Fernandez J, Spezi E, Steenbakkers RJHM, Tanadini-Lang S, Thorwarth D, Troost EGC, Upadhaya T, Valentini V, van Dijk LV, van Griethuysen J, van Velden FHP, Whybra P, Richter C, Löck S (2020) The image Biomarker Standardization Initiative: standardized quantitative Radiomics for High-Throughput Image-based phenotyping. Radiology 295(2):328–338. https://doi.org/10.1148/radiol.2020191145Epub 2020 Mar 10. PMID: 32154773; PMCID: PMC7193906