Extraction of CT-radiomics features for classification of hepatocellular carcinoma (HCC) based on tumour heterogeneity - A pilot study

Bappah Suleiman Yahaya[1,2], Noor Diyana Osman[1,3*], Shazatul Ezanni Samsudin[1,3], Nasibah Mohamad[4], Noor Suzana Ghazali[4], Mohamad Zulfadhli Abdullah[5], Noor Hasyima Mat Zain[5]

[1]Department of Biomedical Imaging, Advanced Medical and Dental Institute, Universiti Sains Malaysia, 13200, Bertam, Penang, Malaysia

[2]Department of Radiography and Radiation Sciences, Federal University of Health Sciences Azare, 751101 Azare, Nigeria

[3]Advanced Management of Liver Malignancies Research Program, Advanced Medical and Dental Institute, Universiti Sains Malaysia, 13200 Kepala Batas, Penang, Malaysia.

[4]Department of Radiology, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia

[5]Imaging Unit, USM Bertam Medical Center, Advanced Medical and Dental Institute, Universiti Sains Malaysia, 13200 Kepala Batas, Penang, Malaysia

*Corresponding author
Noor Diyana Osman
noordiyana@usm.my

Abstract – Radiomics is an emerging field offering the non-invasive extraction of large amounts of quantitative data from CT images, often beyond what the human eye can see. These data, derived from statistical analysis of pixel/voxel intensity and spatial relationships, can aid automated diagnosis, such as LI-RADS classification, with high accuracy. This study aims to extract and analyse first-order CT radiomics features from multiphasic contrast-enhanced liver CT scans to assess their potential for classifying HCC and normal tissues based on tumour heterogeneity. Human ethical approval was obtained from the institutional review board (Approval Code: USM/JEPeM/PP/ 24090793) prior to data collection. This retrospective study extracted and analysed first order radiomics features from contrast-enhanced CT images of radiologically confirmed HCC patients (n = 45) and normal patients (n = 15) acquired between January 2020 and December 2024. Tumour segmentation and radiomics feature extraction were performed using 3D Slicer software (Version 5.8.0), by an expert radiologist. Four core extracted feature metrics including entropy, kurtosis, skewness and variance were evaluated due to their common correlation with tumour heterogeneity. Welch's t-test was used to evaluate statistical differences between the extracted features of HCC and normal tissues. Pixel intensity values were higher in HCC lesions for entropy kurtosis and variance compared to normal liver, indicating increased intensity heterogeneity, asymmetry, and dispersion. Among the analysed metrics, entropy and variance showed the most significant difference between HCC and normal liver tissue ($p \leq 0.0001$). The significant differences in first-order CT radiomics features between HCC and normal liver tissues demonstrate the potential of radiomics in quantitatively characterising tumour heterogeneity. These findings provide a valid foundation for developing machine learning-based classification models for objective and automated LI-RADS classification of HCC.

Keywords – Radiomics, Machine learning, Hepatocellular carcinoma, Classification, LI-RADS, Feature extraction

1 INTRODUCTION

Hepatocellular carcinoma (HCC) is the most common primary liver cancer and a leading cause of global cancer-related deaths [1,2]. The incidence of HCC is greatly high in geographical regions with high prevalence of hepatitis B and C infections, particularly sub-Saharan Africa and Southeast Asia [3].

Its risk factors extend beyond viral hepatitis to include chronic alcohol consumption, non-alcoholic fatty liver disease (NAFLD), exposure to aflatoxin, and metabolic conditions [4,5]. Due to its insidious onset and the lack of early symptoms, diagnosis is mostly at advanced stages, limiting the effectiveness of treatments like surgical resection or liver transplantation, making it one of the leading causes of cancer-related mortality

worldwide. Early detection and accurate characterisation of HCC are crucial for effective management, but poses unique diagnostic and therapeutic challenges [6,7].

Recent advancements in imaging techniques along with enhanced biomarker discovery through feature extraction, have improved early detection rates and allowed for more accurate staging of the disease [8] paving the way toward personalised medicine. Computed tomography (CT) remains the first-line imaging modality for evaluating focal liver lesions due to its wide availability and ability to demonstrate the characteristic enhancement patterns of HCC across multiphasic acquisitions [9,10].

Although conventional radiological assessment relies on subjective visual interpretation, traditional systems such as the Liver Imaging Reporting and Data System (LI-RADS) may be limited in distinguishing subtle textural differences between malignant and normal liver tissues [11]. LI-RADS is a standardised classification system used for characterising liver lesions in patients with cirrhosis or chronic hepatitis B virus (HBV) infection without cirrhosis, groups known to have an elevated risk of HCC. Through continued research into imaging and molecular pathways towards personalised medicine, disease management and outcomes can be enhanced.

Radiomics is a rapidly evolving field in medical imaging that enables the quantitative extraction of imaging features capable of capturing tissue heterogeneity beyond human visual perception [12,13]. These features can reflect the biological characteristics, such as angiogenesis and fibrosis, that are key indicators of tumour behaviour [14]. In HCC, tumour heterogeneity plays a critical role in influencing tumour progression and therapeutic response.

This pilot study aims to extract and analyse CT radiomics features from multiphasic liver CT scans of radiologically confirmed HCC cases and normal liver tissues. The goal is to assess the discriminative potential of first-order statistical features in characterising tumour heterogeneity and provide preliminary evidence for the development of machine learning (ML)-based classification models.

2 MATERIALS & METHODS

2.1 Study population

This retrospective study analysed multiphasic contrast-enhanced CT scans from patients with radiologically confirmed HCC (n = 45) and normal tissue controls (n = 15), collected between January 2020 and December 2024. Ethical approval was obtained from the Institutional Review Board (approval code: USM/JEPeM/PP/24090793). Inclusion criteria included adult patients classified as LI-RADS 4 or 5, with adequate image quality and complete arterial and portal venous phase datasets. The normal control group consisted of individuals with no evidence of liver pathology.

2.2 Image Acquisition and Processing

CT images were acquired using a standard multiphasic liver protocol consisting of non-contrast, arterial, portal venous, and delayed phases. All examinations were performed on a Siemens SOMATOM Definition AS+ CT scanner at the Radiology Department, Hospital Pakar Universiti Sains Malaysia (HPUSM), Kelantan, Malaysia. The CT image sequences were retrieved from the PACS, and the corresponding 3D DICOM volumes were imported into 3D Slicer (Version 5.8.0) software for volumetric segmentation and radiomics feature extraction.

2.3 Segmentation and Feature Extraction

Regions of interest (ROIs) corresponding to HCC tumours were manually delineated by an experienced radiologist using the level-tracing tool on delayed-phase CT volumes in 3D Slicer. Corresponding ROIs of normal liver parenchyma were concurrently segmented for comparative analysis. Radiomics feature extraction was performed using the PyRadiomics extension in 3D Slicer, following standardised preprocessing steps that included voxel resampling and intensity normalisation (Figure 1). A total of 18 first-order statistical features were extracted to characterise the voxel intensity distribution within each ROI. Among these features, four-entropy, kurtosis, skewness, and variance-were selected for detailed evaluation due to their established association with tumour heterogeneity.

2.4 Statistical Analysis

Extracted radiomics features were analysed using Python-based libraries, including scikit-learn and NumPy. Welch's t-test was employed to compare the mean feature values between HCC and normal liver tissues, with a p-value ≤ 0.05 considered statistically significant.

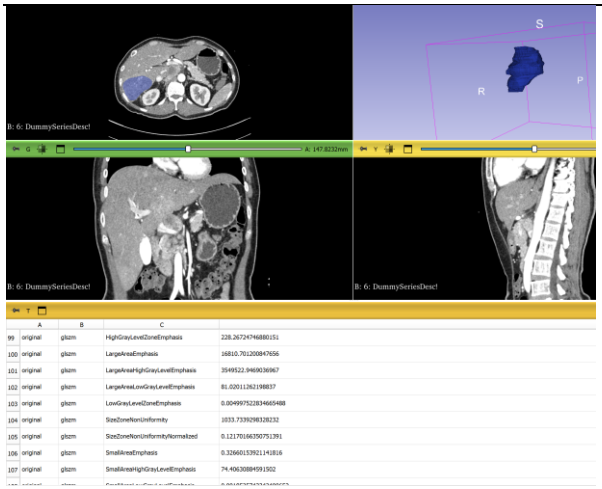

Figure 1: 3D Slicer interface demonstrating ROI segmentation and extracted radiomics features

3 RESULTS & DISCUSSION

A total of 18 first-order radiomics features were extracted from the segmented CT volumes, providing a quantitative assessment of pixel intensity distributions within each ROI. Among these, entropy, kurtosis, skewness and variance were selected for further analysis due to their established relevance to tumour heterogeneity and their strong discriminatory capability between malignant and normal tissues. The inclusion of these four metrics was justified by their sensitivity to subtle pathological alterations in HCC and their proven effectiveness in distinguishing tumour tissue from normal parenchyma. Figure 1 illustrates the distributional differences of these first-order features between the HCC and normal liver groups, further demonstrating their potential utility for quantitative tumour characterisation and classification.

HCC lesions exhibited notably higher entropy values, reflecting increased textural irregularity and greater structural complexity compared with normal liver tissue. Skewness values extended across both positive and negative axes, indicating asymmetric intensity distributions that are likely influenced by necrotic, haemorrhagic or fibrotic components within the tumour. Kurtosis values in HCC were predominantly leptokurtic (kurtosis > 3), in contrast to the near-mesokurtic profile of normal liver tissue (kurtosis ≈ 3). This suggests sharper intensity peaks and greater deviation from a normal distribution, consistent with increased tissue heterogeneity. Among all metrics, entropy and variance demonstrated the most significant differences between HCC and normal groups ($p \leq 0.0001$), indicating strong discriminatory potential.

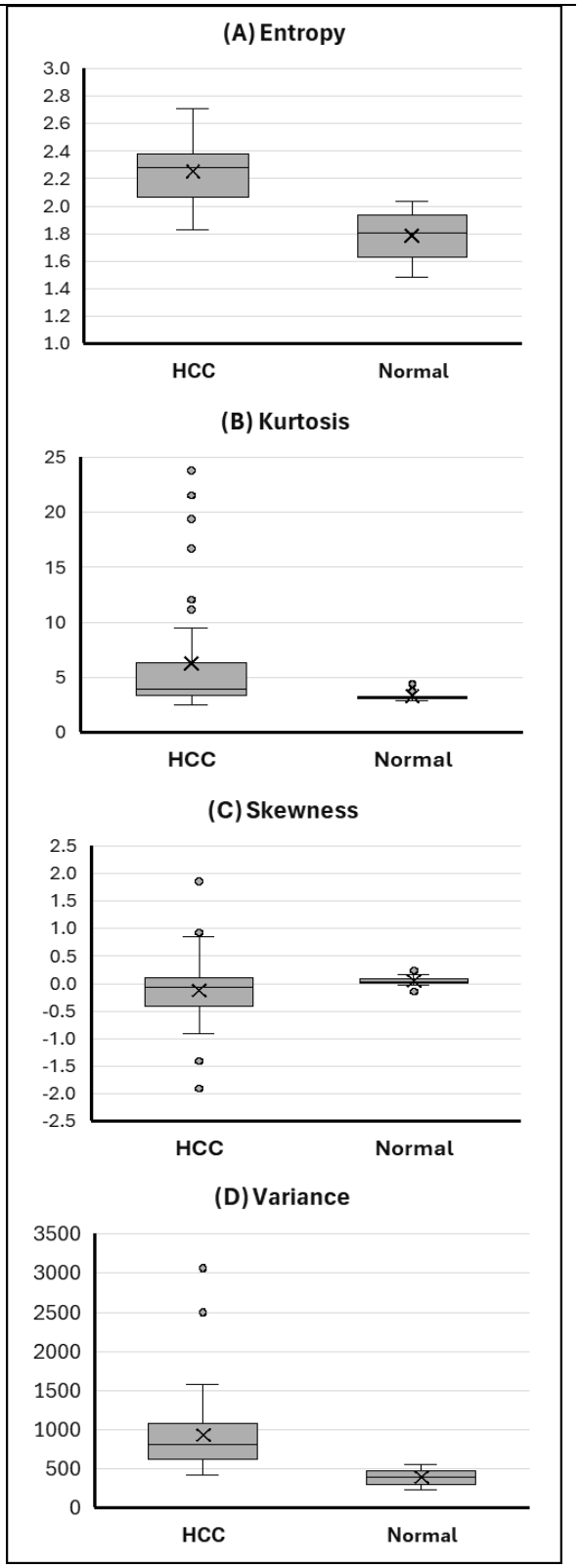

Figure 2: Plots of the comparison of selected first-order radiomics features including (A) entropy, (B) kurtosis, (C) skewness and (D) variance between HCC and normal groups.

The findings of this study are consistent with previous radiomics studies across various malignancies, where entropy and variance have been identified as key markers of tumour grade and aggressiveness [15,17–19]. These radiomic metrics quantitatively capture intra-tumoral heterogeneity through a non-invasive approach, providing the basis for objective differentiation between HCC and normal liver tissue. The analysed feature differences provide the basis for supervised machine learning to automate HCC classification by LI-RADS. Integrating radiomics and ML improves diagnostic accuracy and reproducibility. For personalised cancer care, first-order radiomics features are necessary for robust model training. A limitation of this study is its small sample size and single-centre design. Future work is necessary to build robust classification models, requiring the inclusion of large multicentre datasets, broader radiomics feature classes, and advanced feature selection methods.

4 CONCLUSION

Significant differences in first-order CT radiomics features were observed between HCC and normal liver tissues, particularly in variance and entropy. These findings demonstrate the potential of first-order statistical radiomics features for quantitatively characterising tumour heterogeneity and provide a strong basis for the development of machine learning-based diagnostic tools for automated HCC classification.

ACKNOWLEDGEMENT

The authors acknowledge the support from Imaging Unit, USM Bertam Medical Center (PPUSMB), Penang, for facilitating data collection and analysis.

REFERENCES

[1] Singh, S. P., Madke, T., & Chand, P. (2025). Global epidemiology of hepatocellular carcinoma. Journal of Clinical and Experimental Hepatology, 15(2), 102446.

[2] Zheng, J., Wang, S., Xia, L., Sun, Z., Chan, K. M., Bernards, et al., (2025). Hepatocellular carcinoma: signaling pathways and therapeutic advances. Signal transduction and targeted therapy, 10(1), 35.

[3] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer journal for clinicians, 71(3), 209-249.

[4] McGlynn, K. A., Petrick, J. L., & El-Serag, H. B. (2021). Epidemiology of hepatocellular carcinoma. Hepatology, 73, 4-13.

[5] Samant, H., Amiri, H. S., & Zibari, G. B. (2021). Addressing the worldwide hepatocellular carcinoma: epidemiology, prevention and management. Journal of gastrointestinal oncology, 12(Suppl 2), S361.

[6] Mehta, N. (2020). Hepatocellular carcinoma—how to determine therapeutic options. Hepatology Communications, 4(3), 342-354.

[7] Harding, J. J., Nandakumar, S., Armenia, J., Khalil, D. N., Albano, M., Ly, M., et al., (2019). Prospective genotyping of hepatocellular carcinoma: clinical implications of next-generation sequencing for matching patients to targeted and immune therapies. Clinical Cancer Research, 25(7), 2116-2126.

[8] Lee, S., Choe, E. K., Kim, S. Y., Kim, H. S., Park, K. J., & Kim, D. (2020). Liver imaging features by convolutional neural network to predict the metachronous liver metastasis in stage I-III colorectal cancer patients based on preoperative abdominal CT scan. BMC bioinformatics, 21(Suppl 13), 382.

[9] Shi, Z. X., Li, C. F., Zhao, L. F., Sun, Z. Q., Cui, L. M., Xin, Y. J., et al., (2024). Computed tomography radiomic features and clinical factors predicting the response to first transarterial chemoembolization in intermediate-stage hepatocellular carcinoma. Hepatobiliary & Pancreatic Diseases International, 23(4), 361-369.

[10] Yahaya, B. S., Osman, N. D., Karim, N. K. A., Appalanaido, G. K., & Isa, I. S. (2025). Radiomics and deep learning characterisation of liver malignancies in CT images–A systematic review. Computers in Biology and Medicine, 194, 110491.

[11] Cunha, G. M., Sirlin, C. B., & Fowler, K. J. (2021). Imaging diagnosis of hepatocellular carcinoma: LI-RADS. Chinese Clinical Oncology, 10(1), 3-3.

[12] Zossou, V. B. S., Gnangnon, F. H. R., Biaou, O., de Vathaire, F., Allodji, R. S., & Ezin, E. C. (2024). Radiomics-based classification of tumor and healthy liver on computed tomography images. Cancers, 16(6), 1158.

[13] Jiang, Z. Y., Qi, L. S., Li, J. T., Cui, N., Li, W., Liu, W., & Wang, K. Z. (2022). Radiomics: status quo and future challenges. Artificial Intelligence in Medical Imaging, 3(4), 87-96.

[14] Dreher, C., Linde, P., Boda-Heggemann, J., & Baessler, B. (2020). Radiomics for liver tumours. Strahlentherapie und Onkologie, 196(10), 888-899.

[15] Chen, W., Zhang, T., Xu, L., Zhao, L., Liu, H., Gu, L. R., et al., (2021). Radiomics analysis of contrast-enhanced CT for hepatocellular carcinoma grading. Frontiers in Oncology, 11, 660509.

[16] Haghshomar, M., Rodrigues, D., Kalyan, A., Velichko, Y., & Borhani, A. (2024). Leveraging radiomics and AI for precision diagnosis and prognostication of liver malignancies. Frontiers in Oncology, 14, 1362737.

[17] Zhao, X., Liang, P., Yong, L., Jia, Y., & Gao, J. (2022). Radiomics study for differentiating focal hepatic lesions based on unenhanced CT images. Frontiers in Oncology, 12, 650797.

[18] Tharmaseelan, H., Vellala, A. K., Hertel, A., Tollens, F., Rotkopf, L. T., Rink, J.,et al., (2023). Tumor classification of gastrointestinal liver metastases using CT-based radiomics and deep learning. Cancer Imaging, 23(1), 95.

[19] Hu, S., Lyu, X., Li, W., Cui, X., Liu, Q., Xu, X., ... & Yin, Y. (2022). Radiomics analysis on non-contrast CT for distinguishing hepatic hemangioma (HH) and hepatocellular carcinoma (HCC). Contrast Media & Molecular Imaging, 2022(1), 7693631.