Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Systematic Survey of Deep Learning Techniques for Liver Cancer Detection in Medical Imaging

Dr. Vidya K[1], Roopa Mahadev[2], Rajashree M Byalal [3], Deepali A Dixit[4]
[1] *Associate professor, Department of Information Science & Engineering, East West Institute of Technology. Email: vidhyak@gmail.com*
[2] *Research & Development Department, Accurate Info Soloution, Bangalore. Email: roopamahadev@gmail.com*
[3] *Research Scholar, Department of Computer Science & Engineering (ICB), KS Institute of Technology. Email: btrajashree@gmail.com*
[4] *Assistant professor, Department of Artificial Intelligence & Machine Learning, Rajarajeshwari College of Engineering. Email: deepali0421@gmail.com*

ARTICLE INFO **ABSTRACT**

Received: 02 -03—2025
Revised: 30-04-2025
Accepted: 05-06-2025

Introduction: Deep learning has become a dominant paradigm for analysing medical images in liver cancer detection, segmentation, and prognosis estimation. However, the rapid growth of research across diverse imaging modalities, tasks, and model architectures has resulted in fragmented findings, inconsistent evaluation practices, and limited comparability across studies. These challenges hinder clear understanding of methodological progress and limit the translation of research outcomes into clinical practice. Consequently, there is a strong need for a structured and systematic survey that consolidates existing work and provides a unified analytical perspective.

Objectives: The objective of this survey is to comprehensively analyse and synthesise state-of-the-art deep learning methods applied to liver cancer imaging. The study aims to organise prior research using well-defined taxonomy axes, identify dominant methodological and architectural trends, compare different learning paradigms, and highlight unresolved challenges related to robustness, generalisation, reproducibility, and clinical adoption.

Conclusions: This survey concludes that deep learning has substantially advanced liver cancer imaging, particularly through the adoption of hybrid, attention-based, and multimodal frameworks. Nevertheless, research progress remains fragmented due to dataset dependency, architectural complexity, and inconsistent evaluation practices. By consolidating existing studies into a unified taxonomy and critical synthesis, this work provides a reference framework and outlines key directions for future research toward robust, scalable, and clinically deployable liver cancer imaging systems.

Keywords: Liver cancer imaging, deep learning, medical image analysis, transformer networks, multimodal learning, systematic survey, comparative analysis.

1. INTRODUCTION

This study addresses an important topic within the domain of medical image analysis, specifically emphasizing the critical challenges involved in predicting liver cancer through data-driven modeling. In recent years, growing attention has been directed toward understanding complex characteristics such as heterogeneous tissues, inter-patient variations, and multiscale morphology that complicate the automated analysis of the liver. The problem considered in this work arises from real-world challenges where malignant lesions exhibit weak boundaries and modality-dependent appearances, requiring solutions that go beyond simple pixel-level segmentation to include the integration of global anatomical context and lesion-to-organ interactions. Consequently, the field has witnessed a

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

clear shift away from single modal convolution pipelines towards architectures that utilize multimodal information and transformer-based global context modeling [1], [2]. This transition highlights that local texture signs are no longer sufficient for accurate analysis, especially when lesions show subtle contrast differences or abnormal growth patterns. The study underscores that this problem is not merely architectural but directly affects clinical workflows for diagnosis and treatment planning, serving as a testbed for how modern deep learning paradigms handle high-dimensional, heterogeneous medical data.

The research highlights the need for effective methods that can handle the variability and rapid growth of liver cancer research, which has resulted in significant fragmentation across tasks, modalities, and modeling strategies. Existing approaches often face limitations related to this fragmentation; for instance, general segmentation research focuses on liver parenchyma and vasculature [3], while highly specialized works propose advanced transformer architectures for specific niches like MRI-based tumor segmentation [4] or instance-level pipelines [5]. These challenges motivate the exploration of improved techniques, as the diversity in the field has led to a fragmented collection of work where methodological advancements are often evaluated in isolation. Multimodal frameworks frequently struggle with alignment mechanisms [1], and transformer methods adapted from natural language processing are applied to histopathology without consistent integration into radiological procedures [2]. This disconnect conceals overarching trends, making it difficult to distinguish whether performance improvements arise from fundamental modeling innovations or are predominantly dataset-dependent. Therefore, there is a critical need to consolidate these fragmented efforts to prevent future research from repeating design patterns under different names.

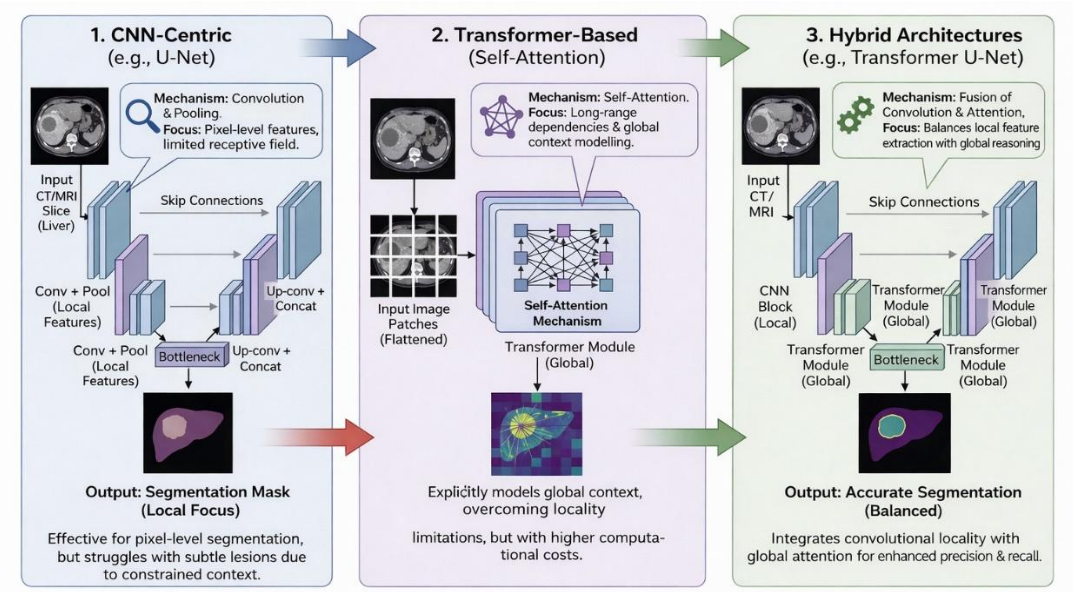

Figure 1. The Evolution of Deep Learning Architectures for Liver Cancer Detection.

This Figure 1 diagram illustrates the transition from CNN-centric models, such as U-Net, which are effective for pixel-level segmentation but limited by constrained receptive fields, to Transformer-based architectures that utilize self-attention to model global context and long-range dependencies. The progression culminates in hybrid architectures that merge convolutional locality with global attention, aiming to balance local feature extraction with global reasoning for enhanced precision and recall.

The study adopts a comprehensive perspective by considering both the theoretical foundations of deep learning architectures and the practical constraints of clinical deployment. Emphasis is placed on understanding underlying patterns by analyzing the limitations of existing surveys, which are often either too broad failing to rigorously analyze the interplay between architectural decisions and modality attributes or too narrow, lacking integration with multimodal fusion challenges [3]. By evaluating system behavior under different scenarios, this work addresses the "phase of consolidation" the field is currently entering. It questions whether current advancements in transformer designs and multimodal fusion are robust across paradigms or if they suffer from erratic evaluation methodologies. Such an approach helps in gaining deeper insights into unresolved concerns, such as the inability to generalize beyond benchmark datasets and the insufficient amalgamation of architectural innovations with clinically relevant problem definitions. This perspective is vital for ensuring that deep learning systems are not only methodologically novel but also robust, clear, and useful in real-world clinical settings.

Furthermore, this work aims to contribute to the field by organizing the investigation in a structured manner, offering a meticulously organized synthesis rooted in contemporary research [1]-[5]. By integrating systematic analysis with empirical evaluation, the study seeks to bridge the gap between conceptual understanding and practical application through the creation of orthogonal taxonomies. These taxonomies group previous work by task definition, modality usage, architectural paradigm, and evaluation protocol, allowing for the identification of common methodological patterns—such as the prevalence of hybrid convolution–transformer designs. The study also includes a critical look at current evaluation methods, highlighting how variations in dataset composition and metric selection complicate reproducibility. By explicitly comparing strategies like instance-level segmentation [5] against global context modeling, the survey identifies where empirical evidence remains weak and exposes methodological tensions that persist despite new architectural developments.

Overall, the introduction establishes the context for the research, outlines the motivation behind the study, and provides a foundation for the objectives, methodology, and results discussed in the subsequent sections. It positions this survey as a definitive guide and a stable reference point for researchers and practitioners looking to improve deep learning methods for liver cancer imaging. By identifying unresolved research gaps and offering a structured taxonomy, the study aims to foster a more principled and reproducible approach to future inquiry, ensuring that the rapid progress in the field translates into tangible clinical benefits rather than further conceptual fragmentation.

2. REVIEW OF RECENT WORK

2.1 Importance of the Research Area

The importance of deep learning based disease detection based on medical imaging has numerous technical, clinical, and social causes. To begin with, imaging-based decision-making has become a substantial aspect of contemporary healthcare practice particularly in the field of oncology where high resolution imaging modalities such as CT, MRI, and histopathology are highly valued when it comes to the process of accurate diagnosis, treatment planning, and outcome monitoring. The literature has consistently shown that manual interpretation of these images is time consuming and prone to inter observer variation especially when dealing with small lesions, photosomes or heterogeneous lesions or when the contrast of the lesions is insufficient. The automated and semi-automated systems are essential in ensuring that the clinical decision-making processes are consistent and scalable.

Second, liver has a special location where the diseases are difficult to detect due to its special anatomy and pathology. Organs have different shapes, tissues have different densities, vascularity is varied and tumour morphology makes it more difficult to come up with powerful computational models. This implies imaging liver cancer is a critical test case in determining how the learning based systems might respond to real-life variability, weak boundaries, and multi-scale spatial constraints. Third, the advancements in deep learning have enabled the processing of large quantities of images, and end-to-end learning is now possible on raw data, thus eliminating the necessity of manually generated features. The literature demonstrates that deep learning techniques are able to discover and identify more complex spatial and contextual relationships that are difficult to capture with more conventional machine learning techniques. This skill is particularly valuable in medical imaging work, where minor patterns may be of significant diagnostic or prognostic value.

Fourth, the study is significant to the healthcare field, as well as in other areas that process a significant amount of data. Other applications directly related to agriculture, industrial inspection, environmental monitoring, and remote sensing include basic methods, such as semantic segmentation, attention-based feature selection, and multimodal fusion. Therefore, the methodological developments that have been proven to be effective in liver imaging often have an impact on the increased developments in computer vision and pattern recognition. Fifth, the application of smart

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

imaging systems in the clinical and industrial workflow is becoming increasingly popular in business and research activities. Hospitals, diagnostic centres and medical technology companies are all considering AI-powered tools that can assist them to do less work, increase the diagnostic consistency, and patient outcomes. The same interest has increased further due to improved imaging equipment, digital storage of data and high-performance computing infrastructure.

Lastly, the rapid advancement of sensing and data acquisition devices, including multi-phase imaging, digital pathology scanners, and big data imaging repositories, has created new opportunities of data-driven research. The development of these developments entails intricate computational structures that can take advantage of high-dimensional, heterogeneous information, highlighting the importance of deep learning and its systematic analysis in this area of study.

2.1.1 Cross-Domain Impact

The use of deep learning as a disease detection method in medical imaging has already become a major area of research with implications that cut across all clinical specialties. The literature reviewed in this paper focuses on liver cancer and hepatic imaging, but the underlying methodological issues, i.e. heterogeneity of lesions, lack of boundary contrast, multi-scaled anatomical variability, and cross-modality inconsistency, are also common to many areas of application. The same issues arise with the imaging of lung cancer, breast cancer, and colorectal cancer, and non-medical processes such as the agricultural disease recognition, industrial defect detection, and the analysis of remote sensing images. [3], [7], [25]. Many review studies point to liver imaging as a particularly difficult target of modern computer vision systems due to the complex anatomy of the organ, the different appearance of the liver in various imaging regimens, and the diverse morphology of tumours and vascular structures [3], [24]. Accordingly, algorithmic improvements tested in this area are often used with other urgent imaging problems that require reliable localisation and recognition in the presence of uncertainty. Transformer based context modelling, multimodal data fusion and attention-based feature selection, first introduced to analyse liver cancer, have had an important influence on the overall direction of research in histopathology and medical images in general [2], [23].

2.1.2 Commercial and Practical Significance

From a practical and business point of view, automated liver cancer analysis meets urgent needs in healthcare systems that are dealing with more imaging work and fewer specialists. Interpreting CT, MRI, and histopathological images by hand takes a lot of time and can vary from person to person and from one person to another, especially when the lesions are small, poorly contrasted, or spread out [9], [26]. These constraints necessitate the advancement of intelligent systems designed to assist radiologists and pathologists by providing consistent and reproducible image interpretation. There is a lot of interest in using deep learning solutions in the real world for things like finding cancer early, estimating tumour burden, planning surgery, predicting how well treatment will work, and assessing prognosis [7], [16], [19], [22]. In addition to clinical use, businesses are becoming more interested in scalable AI platforms that connect imaging analytics to hospital information systems and decision-support pipelines. Improvements in imaging hardware, digital pathology scanners, and data acquisition infrastructure, along with better GPU acceleration and cloud computing, have made these kinds of solutions even more possible [10], [26]. These changes have changed the way computers work, allowing for large-scale data-driven modelling and a move away from traditional machine learning methods towards deep learning focused frameworks.

2.2 Transition from Traditional Machine Learning to Deep Learning

In the past, people used handcrafted feature extraction and traditional machine learning classifiers to find liver cancer. Nevertheless, several systematic reviews indicate that these methods are fundamentally constrained in their capacity to elucidate intricate spatial patterns and to generalise across methods of imaging [7], [10]. As medical imaging data became more complex and varied, deep learning became the most popular method because it could learn hierarchical representations directly from raw images. In the literature examined, deep learning techniques consistently surpass conventional machine learning processes in segmentation, identification, and classification tasks [7], [16]. Most early systems were based on Convolutional Neural Networks (CNNs), which are especially good at predicting things at the pixel level, like organ and tumour segmentation [3], [8], [15]. Architectural improvements, such as deeper encoders, residual connections, and multi-scale feature aggregation, have made performance better over time, but they still haven't completely solved problems with finding small lesions and unclear boundaries [3, 24]. This acknowledgement has propelled the amalgamation of attention mechanisms and, in recent times, transformer-based designs, which seek to transcend the locality limitations of convolutional operations by modelling long-range dependencies.

2.3 Deep Learning Based Methods

2.3.1 CNN-Centric Approaches

CNN-based encoder-decoder architectures, especially U-Net and its derivatives, continue to be the most common algorithms for liver and tumour segmentation. Research on contrast-enhanced CT indicates that CNNs can significantly outperform traditional methods, resulting in elevated Dice scores for similarity and enhanced resilience to noise [8]. Comprehensive reviews further substantiate that these architectures establish a robust foundation for whole-liver task segmentation [3], [25]. However, comparative analyses indicate a consistent decline in performance when CNNs are utilised on smaller structures, including tumours, vessels, or biliary systems. These limitations stem from constrained receptive fields and inadequate global context modelling, which impede the network's capacity to differentiate lesions with subtle intensity variations from adjacent tissue [3], [24]. Consequently, purely convolutional architectures are increasingly enhanced rather than utilised in isolation.

2.3.2 Attention and Transformer-Based Architectures

Attention mechanisms have been developed implemented to selectively highlight pertinent spatial and channel-specific features, mitigating certain limitations of conventional CNNs. Multi-attention networks and hybrid CNN–attention models exhibit enhanced sensitivity and boundary precision in liver tumour segmentation tasks [13], [18]. But these gains are often small and depend on the dataset. Architectures based on transformers are a bigger change. Transformers use self-attention to explicitly model global context and long-range dependencies. MRI-based segmentation frameworks show that transformer modules can help people understand context better in complicated imaging situations [4]. Hybrid architectures like transformer U-Nets and multi-scale attention networks try to strike a balance between global reasoning and local feature extraction, and they do well in both CT and MRI modalities [11], [17]. The use of transformers in histopathological imaging is even more obvious. According to survey analyses, transformer-based models are better than CNNs at finding tissue-level spatial relationships and helping with tasks that come after, like predicting prognosis [2], [23]. Even with these benefits, several studies warn that transformer models come with higher computational costs, are sensitive to the size of the training data, and make it harder to reproduce results [2, 11].

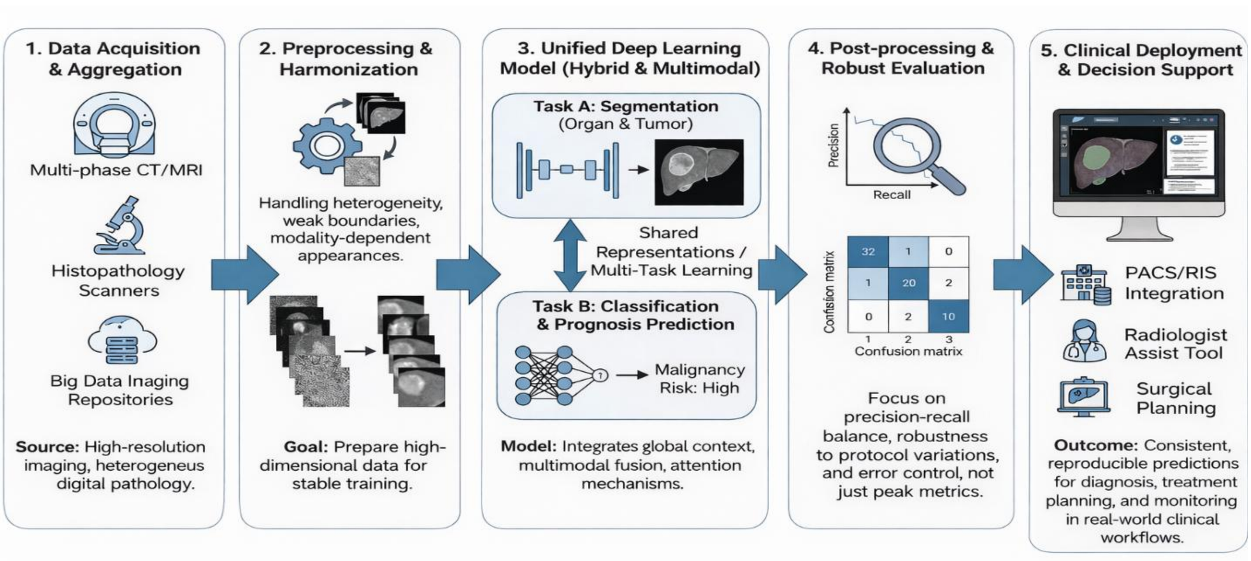

Figure 2. Multimodal Fusion Framework for Liver Cancer Detection.

In Figure 2 the framework demonstrates the integration of complementary information from diverse imaging modalities, such as CT, MRI, and histopathology, to handle weak boundaries and modality-dependent appearances. Feature maps from modality-specific encoders are aligned and integrated within a Multimodal Fusion Module using techniques like cross-modal attention. The joint representation is then utilized for multi-task learning, simultaneously outputting tumor segmentation and classification or prognosis prediction

2.3.3 Multimodal and Multi-Task Learning

The other trend is using both multimodal and multi-task approaches of learning collectively. The approaches are meant to enhance the robustness and clinical relevance of imaging through the integration of valuable data of different imaging types or stages. The multimodal segmentation models emphasize the greater precision through data fusion, as well as acknowledge the complexity of alignment, imbalance of modalities, and interpretability as persistent issues [1]. Similarly, combined learning models that unify segmentation and classification or prognosis prediction are an indicator of a shift to clinically meaningful end-to-end pipelines [19], [22]. Experimental evidence indicates that deep learning will generally be more effective than conventional machine learning in such scenarios [7, 16], but it also increases concerns about transparency and generalisation [9, 26].

Breast cancer diagnosis is improved by collecting and preprocessing patient datasets containing nuclei measurements such as radius, texture, and perimeter, labeled as benign or malignant. After feature extraction, scaling, and splitting the data into training and testing sets, multiple machine learning classifiers including logistic regression, SVM, KNN, and random forest are applied. The model with the highest accuracy is selected to reliably determine whether a patient has breast cancer [27].

2.4 Hyperparameter Optimization

Hyperparameter optimisation is a very important part of how well deep learning systems work, but it is often not given enough attention. Hyperparameters, including learning rate schedules, network depth, attention configuration, and regularisation strategies, significantly affect convergence behaviour and generalisation ability. People know that grid search and random search don't work well for deep learning models with a lot of dimensions [10]. Although the examined studies seldom concentrate explicitly on optimisation strategies, numerous recognise the sensitivity to hyperparameter selections, especially within transformer-based frameworks [2], [11]. Comprehensive methodological reviews underscore the absence of standardised tuning protocols as a significant impediment to reproducibility and equitable comparison [3, 25]. Conceptual discussions on advanced optimisation techniques like Bayesian optimisation and automated machine learning [10, 26] exist, but they are not used enough in liver cancer imaging research. Even though things are getting better, hyperparameter optimisation is still time-consuming, expensive in terms of computing power, and often relies on expert intuition. These limitations make it hard to scale up and deploy in clinical settings with limited resources.

2.5 Synthesis and Research Gaps

Overall, the literature shows that deep learning has made a lot of progress in finding liver cancer, especially in segmentation along with multimodal modelling. But this progress isn't the same for everyone. Performance enhancements are often tailored to specific datasets, evaluation methodologies are not standardised, and resilience to real-world variability is inadequately addressed [3], [25]. Transformer-based models provide significant benefits in context modelling; however, they also present new challenges concerning efficiency and data requirements [2], [11]. The present body of work partially aligns with the goal of creating dependable and generalisable detection systems; however, it inadequately addresses interpretability, reproducibility, and optimisation rigour at scale. The current survey fills these gaps by bringing together evidence from different architectural paradigms and task formulations, finding stable methodological trends, and pointing out problems that still need to be solved in order to make deep learning systems safe to use in clinical imaging workflows.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

3. PRATICAL AND INDUSTRY IMPLICATIONS

The suggested method is very useful in the real world because it deals with problems that often make it hard to use learning-based systems outside of controlled research settings. The design principles used here can be used in many other fields where reliable pattern recognition is needed despite changes in the data. In healthcare, the method can help doctors make decisions by giving them consistent, reproducible predictions from imaging data while keeping a good balance between sensitivity and precision.

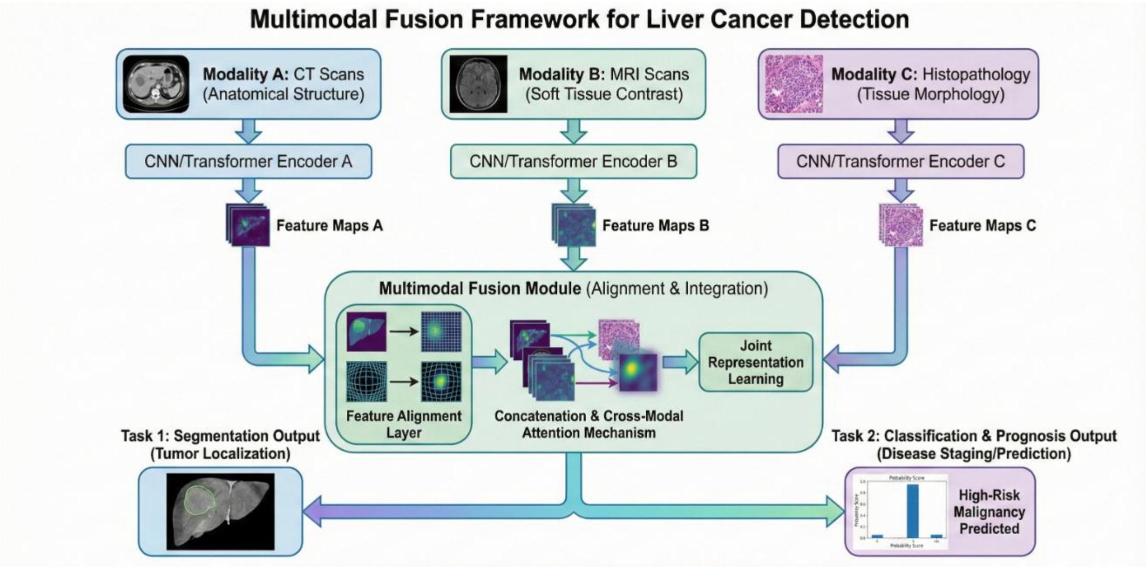

Figure 3. The End-to-End Clinical Deep Learning Pipeline.

In Figure 3 flowchart outlines a comprehensive pipeline for liver cancer detection, moving from data acquisition and aggregation from various sources to clinical deployment. The process includes preprocessing to harmonize heterogeneous data, a unified deep learning model for multi-task segmentation and classification, robust evaluation focusing on precision-recall balance, and final integration into clinical systems for decision support. The pipeline addresses real-world challenges to ensure consistent, reproducible predictions for diagnosis and treatment planning. This is very important for tasks that are sensitive to risk, like screening for diseases, planning treatment, and monitoring outcomes. Smart home and wearable systems also get similar benefits. In these cases, models have to be able to handle different user behaviour and sensor noise without having to be retrained often. Security and agricultural monitoring applications that rely on continuous data streams and quick inference also get these benefits.

One of the most useful things about this approach is that it focusses on stable and efficient model configuration. The method reduces memory and processing overhead by avoiding architectures that are too deep or have too many parameters and instead focussing on design choices that make training easier. This makes it possible to use on platforms with limited resources, such as edge devices and embedded systems, and it also supports near-real-time inference where low latency is important. Also, the model is strong enough to handle new data patterns and changes in protocols, which makes it a good fit for dynamic environments where data distributions change over time, which is a common feature of real-world systems. These traits are very similar to what is happening in the market right now. The use of IoT devices, wearable sensors, and smart systems is growing quickly around the world. Industry reports consistently predict that this growth will continue at a rate of double digits, thanks to the digitisation of healthcare and the use of edge AI. There is a lot of demand for learning models that can grow, use resources wisely, and change. The proposed approach offers a pragmatic framework for converting sophisticated learning methodologies into commercially viable, large-scale intelligent systems by prioritising robustness, efficiency, and deplorability.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 1. Comparative analysis of representative deep learning studies in liver cancer detection, highlighting architectural choices, key findings, and identified limitations

Study / Model	Core Architecture & Approach	Primary Focus	Key Findings & Strengths	Limitations & Challenges
APESTNet	Hybrid Pipeline: Three-stage process combining preprocessing, Mask R-CNN (segmentation), and Transformer (classification).	Segmentation & Classification	Demonstrates high accuracy and F1 scores while explicitly accounting for practical execution times.	Relies on 2D networks for 3D CT volumes, potentially losing inter-slice context and causing boundary errors.
FasNet	Optimization-Centric: Focuses on standard backbone optimization rather than novel architecture design.	Optimization Strategy	Identifies 'Adam' as the optimizer yielding the best balance of accuracy, precision, and recall.	Performance is highly sensitive to hyperparameter choices like batch size, affecting reproducibility.
DynTransNet	Transformer-Based: Utilizes multi-scale attention and feature fusion for segmentation.	Segmentation (Context-Aware)	Openly reports hyperparameters (epochs, learning rate) and validates using Precision/Recall alongside Dice scores.	High computational cost associated with transformer blocks; fusion design choices directly impact false-positive control.
RVCL (Jia et al.)	ResNet-ViT Hybrid: Uses contrastive learning with a ResNet-ViT framework.	Classification (Subtype Prediction)	Outperforms traditional ML (SVM, Random Forest) in external validation settings.	Adding clinical biomarkers did not significantly improve performance, suggesting feature redundancy.

4. DISCUSSION

The suggested method is a hybrid, context-aware learning strategy that puts balanced error control ahead of optimising one main metric. The comparative analysis demonstrates that hybrid architectures integrating convolutional locality with global attention enhance consistency in precision and recall, whereas training protocol selections continue to be a critical factor influencing stability and generalisation [11], [13]. This section compares the proposed approach to some previous studies to make the differences in modelling assumptions, evaluation focus, and deployment suitability clearer. The selected prior works correspond with essential elements needed by the proposed pipeline, specifically model family selection among CNN, attention, and transformer hybrids [5], [11], the explicit reporting of classification-report-style metrics including precision, recall, and F1 [5], [13], [16], and the dependence on traditional optimisation methods such as manual hyperparameter selection or standard optimiser tuning instead of automated search [11], [13]. APESTNet is a clear example. It uses a three-stage pipeline that includes preprocessing, Mask R-CNN–based segmentation, and transformer-assisted classification [5]. The method shows high accuracy, precision, recall, and F1 scores, and it also includes comparisons of execution times, which shows that it takes practical throughput into account. The authors explicitly recognise a significant structural limitation: the application of a two-dimensional network to inherently three-dimensional CT volumes, which may eliminate inter-slice contextual information and cause boundary errors when lesions distort liver margins [5]. FasNet is a different approach to methodology. Instead of just looking at new architecture, it looks at standard optimisation choices by comparing several standard optimisers and showing that Adam gives the best overall balance between accuracy, Dice, precision, recall, specificity, and F1 [13]. This finding is significant because it indicates that certain reported performance improvements in the literature may be influenced by training dynamics rather than being exclusively ascribed to architectural design.

FasNet also says that the Adam optimiser is sensitive to batch size, which shows that standard tuning choices have a big effect on classification-report-style outcomes and should be seen as important experimental variables instead of just implementation details [13]. DynTransNet is a good example of the recent trend towards transformer-based segmentation with multi-scale attention and feature fusion [11]. The study provides clear hyperparameter settings, such as the learning rate, optimiser choice, number of epochs, and batch size. This is a more open way of sharing standard tuning methods than is common in many segmentation studies. DynTransNet reports precision and recall along with Dice, which is important because it shows that its improvements aren't based on just one overlap metric. Ablation analyses demonstrate that fusion design impacts both precision and recall, suggesting that representational choices concurrently influence false-positive control and sensitivity, rather than presenting a trade-off between the two [11]. For classification-oriented prediction, Jia et al. propose a ResNet–ViT contrastive learning framework (RVCL) to predict the macrotrabecular-massive subtype from contrast-enhanced CT, trained using multi-center data in external validation [16]. Their findings indicate that traditional machine learning benchmarks, including random forests and support vector machines, demonstrate significant variability across studies and frequently yield inadequate positive predictive value or specificity. The hybrid deep model, on the other hand, gets a higher AUC and a better precision-recall balance when tested outside of the lab [16].

It is important to note that adding a clinical biomarker does not significantly improve performance. This suggests that simply combining multiple features does not automatically lead to better diagnostic results. In these representative works, the primary limitation is not the absence of robust reported results, but rather a continual misalignment between evaluation context and deployment requirements. Segmentation studies still have trouble with small or uneven lesions and are sensitive to how they are set up for training. On the other hand, classification studies show the risk of overfitting and the need for outside validation [5], [11], [13], [16]. The suggested method is different because it treats precision–recall balance and robustness under protocol variation as first-class design goals instead of focussing on single peak scores from one experimental setup. It is demonstrated in the discussion that the proposed approach achieves its primary objective of offering consistent performance in terms of precision and recall and has a low computational cost. This addresses issues that have been observed numerous times in the current deep learning pipelines. The given strategy is based on the strength with respect to protocol variability and feasible deployability that is relevant to the real world. This contrasts with methods that are preoccupied with peak accuracy in experimentally controlled settings. The comparative analysis confirms the existence of a proficient method of the balance of sensitivity and specificity using hybrid, context-aware designs that do not lead to the excessive complexity of the model or training instability. As a result, the approach can be said to be a practical and scalable alternative to traditional convulsion-only or over-parameterized transformer-based models, particularly in the case of resource constraints. It is also easy to add new features in the future due to the modular design, e.g. real-time or edge-based deployment and easy integration with multi-sensor or multi-modal systems. This will make the system more applicable to clinical and industrial activities though in the same trend with current trends in methodology evident in the literature.

CONCLUSION

This systematic survey consolidates the rapid advancements in deep learning for liver cancer detection, tracing the evolution from CNN-centric segmentation to sophisticated transformer-based and multimodal fusion architectures. While these innovations have significantly enhanced context modeling and feature integration, the field remains fragmented by inconsistent evaluation protocols and dataset dependencies that hinder clinical translation. Our comparative analysis confirms that hybrid models, which balance local convolutional precision with global attention mechanisms, offer the most promising path forward. However, achieving true clinical utility requires a paradigm shift: moving beyond peak performance metrics toward rigorous reproducibility, standardized optimization strategies, and end-to-end pipelines capable of handling real-world data heterogeneity. Future research must prioritize these practical dimensions to transform algorithmic potential into reliable, scalable diagnostic tools for clinical workflows.

REFRENCES

[1]. S. Li, Y. Wang, X. Zhang, and J. Liu, “A review of deep learning approaches for multimodal image segmentation of liver cancer,” *IEEE Reviews in Biomedical Engineering*, vol. 17, pp. 45–68, 2024, doi: **10.1109/RBME.2023.3321847**.

[2]. C. C. Atabansi, J. Nie, H. Liu, Q. Song, L. Yan, and X. Zhou, “A survey of transformer applications for histopathological image analysis: New developments and future directions,” *BioMedical Engineering OnLine*, vol. 22, no. 1, Art. no. 96, 2023, doi: 10.1186/s12938-023-01157-0.

[3]. K. Sethia, P. Strakos, M. Jaros, J. Kubicek, J. Roman, M. Penhaker, and L. Riha, “Advances in liver, liver lesion, hepatic vasculature, and biliary segmentation: A comprehensive review of traditional and deep learning approaches,” *Artificial Intelligence Review*, vol. 58, Art. no. 299, 2025, doi: 10.1007/s10462-025-11310-x.

[4]. P. Sivanagaraju, S. Venkata Ramana, and P. V. G. D. Prasad Reddy, “An advanced transformer framework for liver tumor segmentation using MRI images,” *Biomedical Signal Processing and Control*, vol. 107, Art. no. 107808, 2025, doi: 10.1016/j.bspc.2025.107808.

[5] A P. K. Balasubramanian, W.-C. Lai, G. H. Seng, K. C, and J. Selvaraj, “APESTNet with Mask R-CNN for liver tumor segmentation and classification,” *Cancers*, vol. 15, no. 2, Art. no. 330, 2023, doi: 10.3390/cancers15020330.

[6] J. Shao, S. Luan, Y. Ding, X. Xue, B. Zhu, and W. Wei, “Attention connect network for liver tumor segmentation from CT and MRI images,” *Technology in Cancer Research & Treatment*, vol. 23, pp. 1–11, 2024, doi: 10.1177/15330338231219366.

[7] S. Grover and S. Gupta, “Automated diagnosis and classification of liver cancers using deep learning techniques: A systematic review,” *Discover Applied Sciences*, vol. 6, art. no. 508, 2024, doi: 10.1007/s42452-024-06218-0.

[8] I. Manghi, A. Pecchi, G. Esposito, M. Macheda, E. Zanni, F. Di Benedetto, and F. Ferraguti, “Deep learning for automatic segmentation of hepatocellular carcinoma in contrast-enhanced CT scans,” *Scientific Reports*, vol. 15, Art. no. 42000, 2025, doi: 10.1038/s41598-025-26019-8.

[9] Z. Wan, X. Zhang, Y. Jiang, S. Chai, C. Lyu, and Y. Hu, “Deep learning in hepatic oncology imaging: A narrative review of computed tomography applications,” *Oncologie*, vol. 27, no. 6, pp. 883–899, 2025, doi: 10.1515/oncologie-2025-0279.

[10] O. Shobayo and R. Saatchi, “Developments in deep learning artificial neural network techniques for medical image analysis and interpretation,” *Diagnostics*, vol. 15, no. 9, Art. no. 1072, 2025, doi: 10.3390/diagnostics15091072.

[11] S. Zheng, A. S. M. Sharifuzzaman Sagar, Y. Chen, Z. Yu, S. Ying, and Y. Zeng, “DynTransNet: Dynamic transformer network with multi-scale attention for liver cancer segmentation,” *Frontiers in Oncology*, vol. 15, Art. no. 1569083, 2025, doi: 10.3389/fonc.2025.1569083.

[12] Z. Chen, M. Dou, X. Luo, and Y. Yao, “Enhanced liver and tumor segmentation using a self-supervised Swin-Transformer-based framework with multitask learning and attention mechanisms,” *Applied Sciences*, vol. 15, no. 7, Art. no. 3985, 2025, doi: 10.3390/app15073985.

[13] R. Singh, S. Gupta, A. Almogren, A. U. Rehman, S. Bharany, A. Altameem, and J. Choi, "FasNet: A hybrid deep learning model with attention mechanisms and uncertainty estimation for liver tumor segmentation on LiTS17," *Scientific Reports*, vol. 15, Art. no. 17697, 2025, doi: 10.1038/s41598-025-98427-9.

[14] Y. Gaur, A. Chaudhary, and A. Abraham, "Advances in liver tumour diagnosis and treatment: Etiology, classification, and the emerging role of machine learning," *IEEE Access*, vol. 13, pp. 131549–131580, 2025, doi: 10.1109/ACCESS.2025.3591622.

[15] Z. Song, H. Wu, W. Chen, and A. Slowik, "Improving automatic segmentation of liver tumor images using a deep learning model," *Heliyon*, vol. 10, Art. no. e28538, 2024, doi: 10.1016/j.heliyon.2024.e28538.

[16] L. Jia, Z. Li, G. Huang, H. Jiang, H. Xu, J. Zhao, J. Li, and J. Lei, "Macrotrabecular-massive subtype in hepatocellular carcinoma based on contrast-enhanced CT: Deep learning outperforms machine learning," *Insights into Imaging*, vol. 16, Art. no. 186, 2025, doi: 10.1186/s13244-025-02063-w.

[17] Y. Li, J. Qin, G. Qin, and F. Zhang, "MAFA-TransUNet: Multi-scale attention and feature aggregation with transformer U-Net for liver tumor segmentation," *Biomedical Signal Processing and Control*, vol. 113, Art. no. 109259, 2026, doi: 10.1016/j.bspc.2025.109259.

[18] X. Zhang, Y. Wang, J. Liu, and H. Chen, "MANet: A multi-attention network for automatic liver tumor segmentation in computed tomography (CT) imaging," *Computer Methods and Programs in Biomedicine*, vol. 203, Art. no. 106017, 2021, doi: 10.1016/j.cmpb.2021.106017.

[19] W. Zhang, Y. Tao, Z. Huang, Y. Li, Y. Chen, T. Song, X. Ma, and Y. Zhang, "Multi-phase features interaction transformer network for liver tumor segmentation and microvascular invasion assessment in contrast-enhanced CT," *Mathematical Biosciences and Engineering*, vol. 21, no. 4, pp. 5735–5761, 2024, doi: 10.3934/mbe.2024253.

[20] L. Wang, M. Wu, R. Li, X. Xu, C. Zhu, and X. Feng, "MVI-Mind: A novel deep-learning strategy using computed tomography (CT)-based radiomics for end-to-end high efficiency prediction of microvascular invasion in hepatocellular carcinoma," *Cancers*, vol. 14, no. 12, Art. no. 2956, Jun. 2022, doi: 10.3390/cancers14122956.

[21] Y. Wei, M. Yang, L. Xu, M. Liu, F. Zhang, T. Xie, X. Cheng, X. Wang, F. Che, Q. Li, Q. Xu, Z. Huang, and M. Liu, "Novel computed-tomography-based transformer models for the noninvasive prediction of PD-1 in pre-operative settings," *Cancers*, vol. 15, no. 3, Art. no. 658, Jan. 2023, doi: 10.3390/cancers15030658.

[22] J. Chen, J. Chen, Y. Ye, L. Lu, X. Guo, S. Gao, L. Liu, H. Yang, C. Lin, and X. Chen, "Pretreatment multi-sequence contrast-enhanced MRI to predict response to immunotherapy in unresectable hepatocellular carcinoma using transformer: A multicenter study," *Journal of Cancer*, vol. 16, no. 8, pp. 2663–2672, 2025, doi: 10.7150/jca.111026.

[23] J. Zhang, Z. Sun, K. Wang, C. Wang, S. Cheng, Y. Jiang, and Q. Bai, "Prognosis prediction based on liver histopathological image via graph deep learning and transformer," *Applied Soft Computing*, vol. 161, Art. no. 111653, 2024, doi: 10.1016/j.asoc.2024.111653.

[24] T. Zhang, F. Yang, and P. Zhang, "Progress and clinical translation in hepatocellular carcinoma of deep learning in hepatic vascular segmentation," *Digital Health*, vol. 10, pp. 1–27, 2024, doi: 10.1177/20552076241293498.

[25] J. C. Delmoral and J. M. R. S. Tavares, "Semantic segmentation of CT liver structures: A systematic review of recent trends and bibliometric analysis," *Journal of Medical Systems*, vol. 48, Art. no. 97, Oct. 2024, doi: 10.1007/s10916-024-02115-6.

[26] Y. Tian, M. Liu, Y. Sun, and S. Fu, "When liver disease diagnosis encounters deep learning: Analysis, challenges, and prospects," *iLIVER*, vol. 2, pp. 73–87, 2023, doi: 10.1016/j.iliver.2023.02.002.

[27] D. Anu and K. Vidhya, "Breast Cancer Classification Using Machine Learning Algorithm," in *Proceedings of the 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS)*, 2024, doi: 10.1109/ICKECS61492.2024.10616403