Abstract

Gallbladder Cancer (GBC) is the most common biliary tract cancer and the 5th most common gastrointestinal tract malignancy. India sees about 20% of annual GBC-related deaths worldwide and faces an incidence rate compared to the global highest. The overall mean survival rate for patients with advanced GBC is only six months, and the 5-year survival rate is less than 5%. Early diagnosis and curative surgical resection remain the only hope to improve the bleak survival statistics. Ultrasound (USG) is a popular and excellent candidate diagnostic modality for abdominal ailments in low-resource settings due to its low cost, availability, and ionizing radiation-free nature. USG is also the first-line diagnostic modality for gallbladder (GB) diseases. However, diagnosing GBC in USG is difficult, even for experienced radiologists, due to the overlapping visual features of benign and malignant GBs and various confounding medical conditions such as cholecystitis, renal failure, pancreatitis, and Rokitansky-Aschoff sinuses. Our experiments reveal that human experts could achieve only about 70% sensitivity (recall) in GBC detection from USG.

Inspired by the recent success and the transformational capabilities shown by Machine Learning (ML) models, especially the Deep Neural Networks (DNNs), in a plethora of medical image computing tasks, we investigate leveraging DNNs to detect GBC from USG. However, the low image quality arising from noise and artifacts such as shadows or textures, the operator bias and variation in view due to handheld sensors, and the lack of annotated data make the application of DNNs difficult in USG.

In our pursuit of enhancing diagnostic capabilities, we systematically explore the potential of deep convolutional models, leading to the development of GBCNet. GBCNet is a two-stage DNN that first localizes the GB or the region-of-interest (ROI), and then employs a specialized classifier based on multi-scale, second-order pooling (MS-SoP) for robust GBC detection. We further develop a Gaussian smoothing-based training curriculum inspired by human visual acuity to mitigate the effect of spurious textures. GBCNet tackles issues such as noise, artifacts,

and viewpoint variation in USG imaging, improving the GBC detection sensitivity by 7 points compared to SOTA DNN models and 20 points compared to expert radiologists.

The reliance on bounding box annotations for training GBCNet's localization component presents a significant bottleneck, given the high cost and complexity of obtaining such annotations. To overcome this challenge, we delve into utilizing limited supervised data by introducing – (1) an unsupervised contrastive framework for learning GBC representations from unlabelled video data, and (2) a weakly supervised object detection technique to use only image-level labels instead of bounding box annotation requirements, thus designing the more practical models for real-world deployment.

We address the crucial aspect of interpretability in GBC detection by introducing RadFormer, a deep neural network architecture capable of generating interpretable explanations for its decisions. RadFormer not only improves detection sensitivity over GBCNet but also aids in understanding the underlying visual features relevant to GBC diagnosis, bridging the gap between AI-based detection and clinical interpretability.

Finally, we advocate for a paradigm shift towards video-based GBC detection, leveraging the rich spatiotemporal information available in full USG videos. Video-based detection is also clinically more relevant as single frames may not contain conclusive evidence for disease detection. We introduce an innovative masked auto-encoder design called FocusMAE, to learn self-supervised representations for GBC from USG videos. We demonstrate significant improvements in GBC detection using FocusMAE, achieving a 100% sensitivity and thus showcasing the potential of video-based approaches in streamlining the detection process and reducing operator-specific variations.

In summary, we designed and developed accurate, data-efficient, interpretable, and clinically relevant DNN models which could overcome the challenges such as noise, artifacts, viewpoint variability, data scarcity, and real-time applicability in detecting GBC from USG, thereby opening future avenues for transformational research.