Advanced Gallbladder Segmentation in Dynamic Ultrasound Imaging Using Fully Convolutional Networks

You-Jie Chen[1], Tai-Been Chen[2,*], Wen-Hung Twan[3]

[1]The Affiliated Senior High School of National Kaohsiung Normal University, Kaohsiung, Taiwan

[2]Department of Radiological Technology Faculty of Medical Technology, Teikyo University, Tokyo, Japan

[3]Department of Life Sciences, National Taitung University, Taitung, Taiwan

Received 28 April 2024; received in revised form 28 June 2024; accepted 01 July 2024

DOI: https://doi.org/10.46604/emsi.2024.13650

Abstract

This study develops an advanced technique for segmenting the gallbladder from dynamic B-mode ultrasound images to enhance the accuracy and efficiency of volumetric analysis in medical diagnostics. Using a Wi-Fi probe, volumetric data are captured and processed into labeled images for training a fully convolutional network (FCN) model with specifications including an epoch of 9, a batch size of 3, and a learning rate of 0.001. Performance metrics such as global accuracy, mean accuracy, and Intersection over Union (IoU) are evaluated. The MobileNetV2 architecture achieves a maximum mean IoU of 0.690 and a mean Boundary F1 (BF) score of 0.990, while the ResNet50 architecture demonstrates significant effectiveness. This study substantiates the effectiveness of the MobileNetV2 architecture for precise gallbladder segmentation in dynamic B-mode ultrasound imaging.

Keywords: FCN, Dynamic B-mode Ultrasound, Wi-Fi Probe, Ultrasound Gallbladder Image

1. Introduction

The gallbladder, a pivotal organ within the human digestive system, is important to various disorders such as gallstones, cholecystitis, and polyps, each carrying substantial health risks. It is fundamental to develop accurate segmented methods for depicting the gallbladder with ultrasound images. Ultrasound is a handy and non-invasive tool with wide availability and cost-efficiency. Meanwhile, ultrasound scanning is also widely used to examine the gallbladder and other soft tissues inside the human body.

However, the quality of the images is often scattering and spike noise due to the sound wave attenuation in depth. Such noise in images was produced by the transducer, moving objects, and the operator skill. Moreover, dynamic ultrasound imaging involves the real-time frames of moving organs and it is difficult to perform image analysis accurately. In general, the segmentation of ultrasound gallbladder was based on manual or semi-automated techniques. Typically, segmentation requires considerable labor hours and expertise in human anatomy. The segmentation of ultrasound is usually using machine learning approaches. Fully convolutional networks (FCNs) have been proven to be an effective method for segmentation of any kind of image. FCN is designed for segmentation of target regions of the image [1, 2]. Therefore, the FCN is the choice to do the segmentation of the gallbladder for the dynamic ultrasound images.

Accurate gallbladder segmentation enhances visualization and characterization, crucial for clinicians to identify abnormalities and assess their size, shape, and other vital diagnostic features [3-10]. Such precision is pivotal in planning both

* Corresponding author. E-mail address: ztbchen@outlook.com

English language proofreader: Hao-Lun Chang

surgical and nonsurgical treatments. Specifically, in surgeries like cholecystectomy (gallbladder removal), accurate segmentation is key for preoperative planning, which renders a detailed anatomical map to guide surgeons during minimally invasive procedures, thereby reducing complications and improving outcomes [11-13]. Additionally, in managing chronic gallbladder conditions, segmentation is indispensable for monitoring disease progression or evaluating treatment efficacy, enabling healthcare professionals to make informed decisions about adjusting or discontinuing treatments [14-16].

The gallbladder segmentation is important to proffer information on location and shape for surgical operation and other applications. The merits of segmentation techniques are an integration of the FCN model for dynamic ultrasound video and help to enhance the accuracy of operation. FCN approach is one of the powerful segmented tools that offer high accuracy and efficiency [17, 18]. The gallbladder segmentation of dynamic ultrasound video has marked an uneasy task. The properties of FCN architecture are to learn efficiency and adaptability features from input data.

The motivation is to adopt FCNs for segmentation of the gallbladder in dynamic ultrasound imaging primarily focusing on enhancing the efficiency and accuracy of the diagnostic process. By automating the segmentation task, FCNs significantly reduce the workload on medical professionals, potentially accelerating diagnostic procedures and enabling the treatment of a larger number of patients. This work trains an FCN model that could automatically segment the gallbladder in a dynamic ultrasound video. All sources and references in this paper have been appropriately credited and have rendered originality.

2. Methodology

The methodology of this research followed a structured path, beginning with the acquisition of ultrasound videos. These videos were transformed into individual static images or frames. There are steps involved in labeling frames with human drawing boundaries of the gallbladder, followed by the training FCN models and obtaining final results. The entire research process is encapsulated in Fig. 1, providing a visual overview of the workflow.

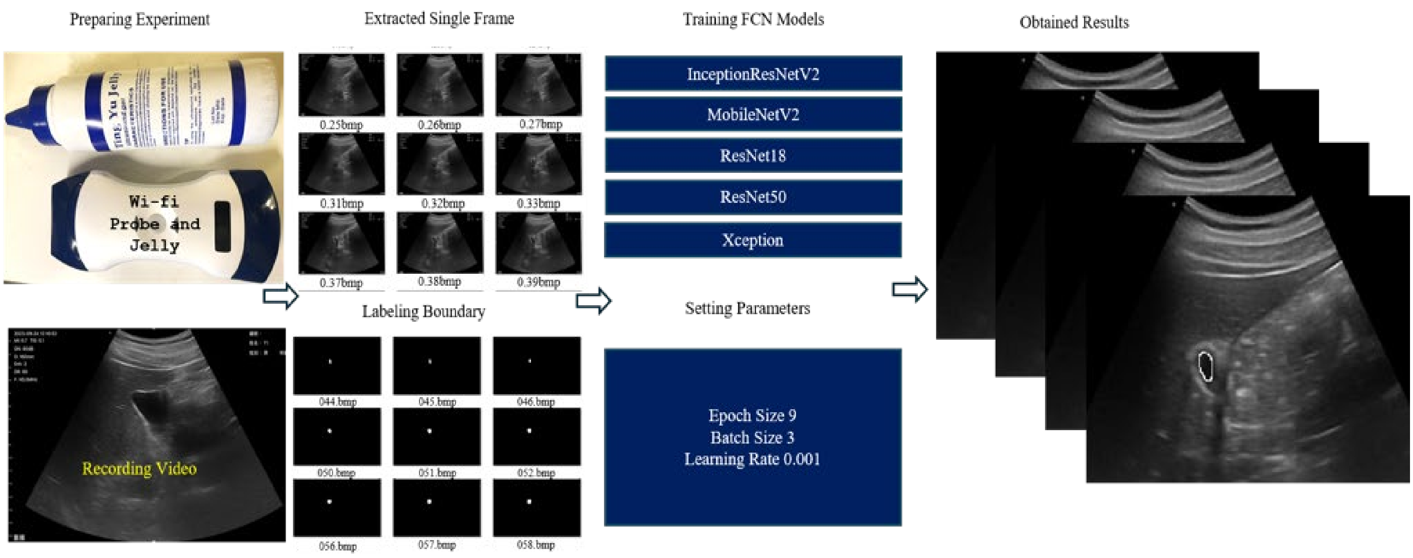


Fig. 1 Diagram illustrating the research methodology

2.1. Data Collection

High-resolution dynamic ultrasound videos of the gallbladder were collected from one of the authors using a cutting-edge Wi-Fi probe and jelly, resulting in detailed volumetric data. These videos, each 11 seconds in length with a frame rate of 9 frames per second and a resolution of 1088×720 pixels, are illustrated in Fig. 2. The videos underwent a systematic process to extract static B-mode ultrasound images, from which forty labeled B-mode sonography images were compiled. This curated dataset was subsequently utilized to train a fully convolutional network (FCN) model.

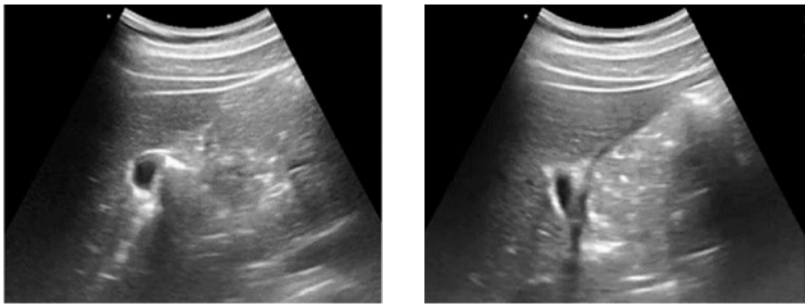


Fig. 2 Displays two frames from the ultrasonic video

2.2. Image Labeling

Experienced radiologists meticulously labeled the extracted images, providing accurate outlines of the gallbladder's boundaries, as depicted in Fig. 3. This careful procedure yielded a dataset comprising forty labeled B-mode sonography images.

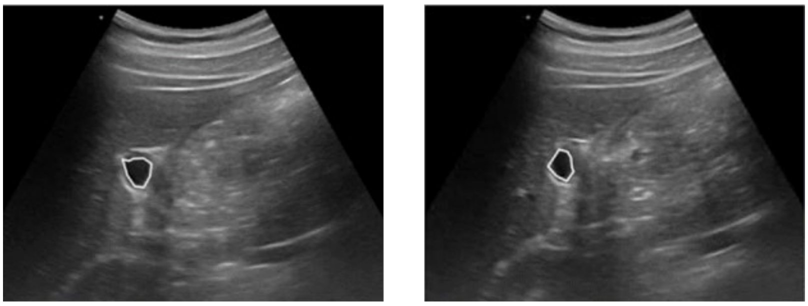


Fig. 3 Left is one frame of ultrasonic video concerning the label of the gallbladder (right)

2.3. Training FCN Model

The FCN model underwent training on a meticulously labeled dataset, utilizing the MobileNetV2 and ResNet50 architectures known for their effectiveness in image segmentation tasks. The training parameters were introduced as an epoch count of 9, a batch size of 3, and a learning rate of 0.001. The models' candidates include InceptionResNetV2, MobileNetV2, ResNet18, ResNet50, and Xception. Using them to train segmentation of gallbladder for dynamic ultrasound video based on their excellent deep learning architecture. Meanwhile, a personal computer is configured as RAM 1 TB, hard disk (SSD) TB, and Nvidia GTX-4090 GPU for training FCN models. These components are particularly adept at navigating the complexity and variability characteristic of medical imagery, as seen in ultrasound scans of the gallbladder. Their depth and complexity enable FCN models to learn detailed and subtle image representations, which are essential for achieving precise segmentation.

2.3.1 Dataset Split and Rationale

In this study, a structured approach was applied to split the dataset into training, validation, and test sets. The dataset consisted of 40 labeled and raw B-mode sonography images. The division was performed as follows: training set 20 images (50%), validation set 10 images (25%) and testing set 10 images (25%). The rationale behind this split was to ensure enough data for training the model while also preserving enough samples to validate and test the model's performance. Given the limited dataset size, maintaining a balance was crucial to avoid overfitting and to ensure the model's ability to generalize.

2.3.2 Data Augmentation to Overcome Potential Limitations

The relatively small dataset size of 40 labeled images presents potential limitations concerning the generalizability of the findings. Small datasets can incur overfitting, where the model learns the specific features of the training data rather than general patterns applicable to new data. This could reduce the model's effectiveness when applied to a broader range of images with different characteristics. In this study, a data augmentation technique was used to simulate twelve different levels of spike noise and contrasts in ultrasound images (Fig. 4). The original image after being normalized as 0 to 1, denoted as Y, underwent a transformation, based on a normal distribution to generate new images (Y^{new}) with varying noise and contrast levels. The process is mathematically represented as Eq. (1).

$$Y^{new} = Y + |\alpha| \tag{1}$$

Where α is a random number sampling from N(0, 1) representing a normally distributed variable with mean 0 and standard deviation 1.

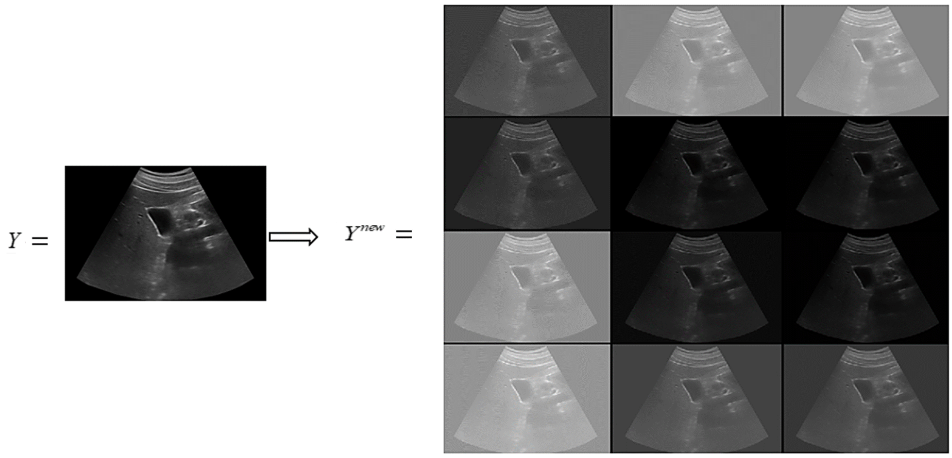
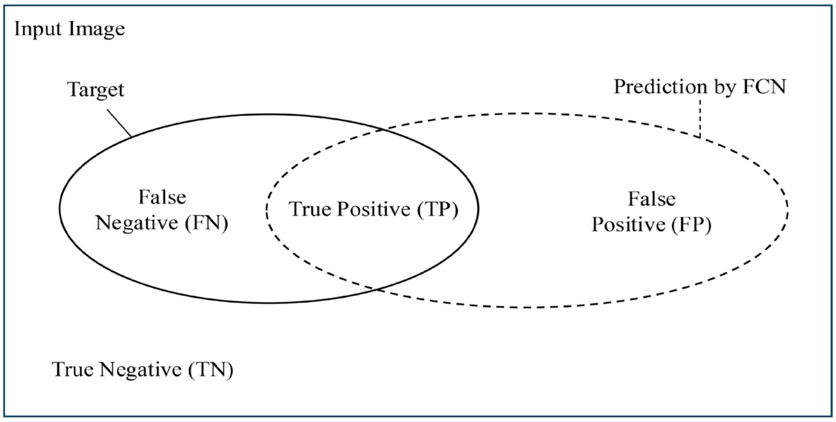


Fig. 4. Data augmentation for ultrasound images using spike noise and contrast variations.

2.4. Evaluation Metrics

The evaluation of the model's performance encompassed a range of metrics, including global accuracy, mean accuracy, mean Intersection over Union (IoU), weighted IoU, and mean Boundary F1 (BF) score, as illustrated in Fig. 5. This suite of metrics offered a thorough evaluation of the model's proficiency in segmentation accuracy and its ability to delineate boundaries accurately.



Global Accuracy = (TP+TN)/(TP + FN + FP+TN)

IoU = Intersection/Union = TP/(TP + FN + FP)

BF score = 2(Recall × Precision)/(Recall + Precision) = 2TP/(2TP + FN + FP)

Recall = TP/(TP + FN) (also call **mean accuracy**, true positive rate or sensitivity)

Precision = TP/(TP + FP) (also call positive predictive value or specificity)

Fig. 5 The evaluated matrix for training the FCN model

3. Results

The FCN methods used in this study are Inceptionresnetv2, MobileNetV2, Resnet18, Resnet50, and Xception. The MobileNetV2 model furnished the highest mean Intersection over Union (IoU) score of 0.690 and a mean Boundary F1 (BF) score of 0.990 among those of the other methods. The Resnet50 architecture also demonstrated substantial effectiveness, recording marginally lower mean IoU and BF scores in comparison (as detailed in Fig. 1 and Fig. 6).

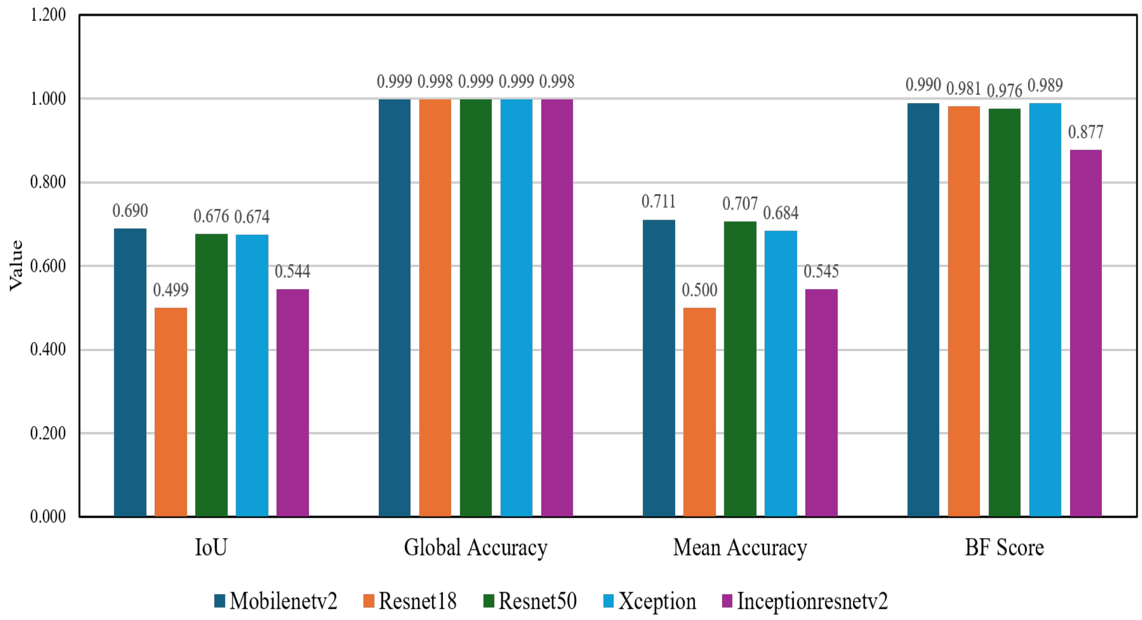


Fig. 6 Performance Metrics Comparison of Five FCN Models.

The graphs and detailed discussion highlight the relative strengths and weaknesses of each model. While MobileNetV2 shows slightly better performance in terms of IoU, BF Score, and Mean Accuracy, the differences, though small, can be crucial in high-stakes medical applications. Therefore, the choice of model should consider the specific requirements of the task at hand, balancing the need for precision with computational efficiency and robustness. The IoU is a critical metric for assessing the accuracy of segmentation models. The difference of 0.014 (1.4%) between MobileNetV2 (0.690) and ResNet50 (0.676) might appear negligible at first glance. However, in applications requiring high precision, such as surgical planning or automated diagnosis, even a small improvement in IoU can elicit better clinical outcomes.

For instance, more accurate segmentation can reduce the risk of missing critical regions, thus improving the reliability of the diagnostic process. The Boundary F1 Score measures how well the model delineates the boundaries of the segmented object. MobileNetV2 achieves a BF Score of 0.990 compared to ResNet50 of 0.976. The 0.014 difference indicates that MobileNetV2 is slightly better at capturing the precise boundaries of the gallbladder, which is crucial in applications where boundary accuracy directly impacts the effectiveness of subsequent medical interventions. MobileNetV2 also shows a higher mean accuracy (0.711) compared to ResNet50 (0.707). Although the difference is small, it suggests that MobileNetV2 has a slight edge in correctly identifying the gallbladder across different images. This can be particularly important in real-world clinical settings where variability in patient anatomy and imaging conditions can affect model performance.

These findings render substantial implications for the field of medical imaging, especially for non-invasive diagnostics. The accuracy and boundary precision achieved by models like MobileNetV2 and ResNet50 are promising for enhancing the diagnosis and treatment planning of gallbladder-related conditions. This comparative analysis highlights the importance of choosing the right FCN models for medical image segmentation, considering their unique strengths and challenges. MobileNetV2, in particular, emerges as a potential benchmark setter in dynamic ultrasound imaging analysis, offering a promising blend of accuracy and efficiency.

Fig. 7 presents a side-by-side comparison of gallbladder segmentation: one side is the segmentation marked by a clinical expert (True Label), and the other side is the segmentation generated by the AI using the MobileNetV2 framework (AI Label). The left image shows the manual segmentation, where a clinician has delineated the gallbladder boundaries, establishing a reference standard or "ground truth." On the right, the AI segmentation was demonstrated, showcasing the outcomes of the automated process. This juxtaposition is designed to illustrate the model's ability to match the precision of the segmentation of an expert. The resemblance between the two images, particularly in the outlined regions, serves as visual evidence of the

AI's proficiency in accurately identifying and outlining the gallbladder's structure in ultrasound images. The performance of the MobileNetV2 model is further indicated by the proximity of its generated boundaries to the true labels. This comparison provides a tangible demonstration of the model's effectiveness in medical image segmentation, highlighting its potential utility in clinical settings.

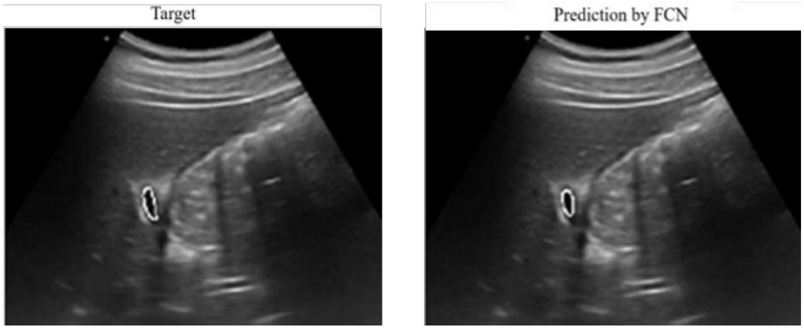


Fig. 7 True (Left) and AI (Right) labels generated by Mobilenetv2 model

4. Discussions

The high mean IoU achieved by MobileNetV2 not only underscores its effectiveness in accurately segmenting the gallbladder but also emphasizes the role of advanced neural network architectures in medical imaging. This distinction in performance, particularly against traditional methods, highlights the transformative potential of deep learning models in clinical applications. The comparative analysis between MobileNetV2 and ResNet50 reveals nuanced insights into how architectural variations can impact feature extraction capabilities, a critical aspect of image segmentation. MobileNetV2's efficiency in handling image data, coupled with its robustness in delineating intricate structures like the gallbladder, suggests its suitability for real-time diagnostic applications.

On the other hand, the slightly lower performance of ResNet50 could be leveraged to understand the trade-offs between complexity and accuracy in model selection. The implementation of FCN-based segmentation, as demonstrated in this study, could be a major step forward in reducing diagnostic errors and enhancing the early detection of gallbladder diseases. This is especially significant considering the challenges posed by the varying shapes and sizes of gallbladders, as well as the dynamic nature of ultrasound imaging. Beyond the scope of gallbladder diagnostics, the success of these models points to a larger trend in medical imaging, i.e., the growing reliance on and integration of advanced machine-learning techniques The presented methods are opening new avenues for exploration and innovation in ultrasound dynamic imaging. However, there are still some limitations in current deep learning approaches. The dependence on large datasets for training, potential biases in data, and the need for extensive computational resources are factors that must be addressed for these technologies to be more widely adopted.

This work also opens discussion on the practical implications of uncooperative patients. Both deep learning and machine learning models are promisingly and pervasively deployed in clinical applications. However, adjusting training hyperparameters is instrumental, especially during creating deep learning and machine learning models. Therefore, the efficacy of FCN-based models like MobileNetV2 in the segmentation of gallbladder with dynamic ultrasound has been successful. Meanwhile, improving the applications can be further investigated in the broader field of medical imaging. The continual evolution of machine learning technologies promises enhanced diagnostic accuracy and efficiency and symbolizes a need for a deeper understanding of their impact on healthcare as a whole.

5. Conclusions

This study substantiates the effectiveness of the MobileNetV2 architecture for precise gallbladder segmentation in dynamic B-mode ultrasound imaging. The presented results demonstrate that MobileNetV2 improves segmentation accuracy and expedites the diagnostic process, which is crucial for prompt and efficient clinical decision-making. The notable performance of MobileNetV2, reflected in high mean Intersection over Union (IoU) and Boundary F1 (BF) scores, marks a significant advancement in medical imaging, where accuracy and speed are of utmost importance. The integration of FCN models, such as MobileNetV2 into clinical practice, offers more than technical benefits; it has the potential to revolutionize patient care by enhancing diagnostic precision. The application of FCNs holds promise for the creation of real-time diagnostic tools, providing healthcare professionals with vital support in the early detection and management of gallbladder conditions.

Additionally, this research lays the groundwork for further investigations into the application of FCNs in a wider range of ultrasound imaging contexts. It opens avenues for exploring how these AI-driven tools can be standardized and implemented across different healthcare settings, ensuring consistent and reliable patient care. While the current findings are encouraging, continuous research and validation within clinical settings remain essential. Enhancing dataset diversity, refining the models, and further evaluating these technologies in practical scenarios are critical steps toward the seamless integration of advanced machine-learning methods into routine medical diagnostics. Such efforts are poised to bring significant benefits to both healthcare professionals and patients.

6. Research Limitations and Future

The training of FCN model on a dataset, comprising forty B-mode sonography images. presents limitations in fully capturing the diverse spectrum of gallbladder pathologies and patient demographic variations. The inclusion of a larger, more varied dataset could significantly enhance the model's ability to generalize across different conditions. Currently, the evaluation of the FCN model has been confined to a controlled environment, leaving its efficacy in varied real-world clinical scenarios, which involve diverse imaging conditions and equipment, unverified.

Future research should aim to expand the dataset to encompass a wider array of gallbladder conditions, include a broader demographic representation, and incorporate images from various types of ultrasound machinery. Such diversification is expected to bolster the model’s robustness and practical utility in clinical settings. The development of automated AI-driven labeling methods could also prove instrumental in minimizing human bias and reducing the workload involved in dataset preparation for FCN training. Furthermore, investigating the integration of FCN-based segmentation into prevailing diagnostic procedures and its consequent impact on clinical decision-making processes would offer significant insights and advancements in the field.

Conflicts of Interest

The authors declare no conflict of interest.

Ethical Approval: The study was exempt from Institutional Review Board (IRB) review as the data was collected from one of the listed authors. Therefore, no formal ethical approval was required.

References

[1] M. Zhuo, X. Chen, J. Guo, Q. Qian, E. Xue, and Z. Chen, “Deep Learning-Based Segmentation and Risk Stratification for Gastrointestinal Stromal Tumors in Transabdominal Ultrasound Imaging,” Journal Ultrasound in Medicine, in press. https://doi.org/10.1002/jum.16489.

[2] X. Jia, X. Li, T. Shen, L. Zhou, G. Yang, F. Wang, et al., “Monitoring of Thermal Lesions in Ultrasound Using Fully Convolutional Neural Networks: A Preclinical Study,” Ultrasonics, vol.130, article no. 106929, April 2023.

[3] B. He, S. Zhao, Y. Dai, J. Wu, H. Luo, J. Guo, et al., “A Robust and Automatic CT-3D Ultrasound Registration Method Based on Segmentation, Context, and Edge Hybrid Metric,” Medical Physics, vol. 50, no. 10, pp. 6243-6258, October 2023.

[4] V. T. Hoang, H. A. T. Van, T. T. T. Nguyen, V. Chansomphou, and C. T. Trinh, “Diffuse Gallbladder Adenomyomatosis with an Inflammatory Complication in an Adult,” Case Reports in Gastroenterology, vol. 15, no. 1, pp. 100-107, January-April 2021.

[5] R. Jin, M. Wang, L. Xu, J. Lu, E. Song, and G. Ma, "Automatic 3D CT Liver Segmentation Based on Fast Global Minimization of Probabilistic Active Contour," Medical Physics, vol. 50, no. 4, pp. 2100-2120, April 2023.

[6] J. Mitra, C. Bhushan, S. Ghose, D. Mills, A. Patel, H. Chan, et al., "A Hybrid Deformable Registration Method to Generate Motion-Compensated 3D Virtual MRI for Fusion with Interventional Real-Time 3D Ultrasound," International Journal of Computer Assisted Radiology and Surgery, vol. 18, no. 8, pp. 1501-1509, August 2023.

[7] Z. Szentimrey, S. de Ribaupierre, A. Fenster, and E. Ukwatta, "Automated 3D U-Net Based Segmentation of Neonatal Cerebral Ventricles from 3D Ultrasound Images," Medical Physics, vol. 49, no. 2, pp. 1034-1046, February 2022.

[8] Y. Chen, L. Xing, L. Yu, W. Liu, B. P. Fahimian, T. Niedermayr, et al., "MR to Ultrasound Image Registration with Segmentation-Based Learning for HDR Prostate Brachytherapy," Medical Physics, vol. 48, no. 6, pp. 3074-3083, June 2021.

[9] M. F. Azampour, M. Tirindelli, J. Lameski, M. Gafencu, E. Tagliabue, E. Fatemizadeh, et al., "Anatomy-Aware Computed Tomography-to-Ultrasound Spine Registration," Medical Physics, vol. 51, no. 3, pp. 2044-2056, March 2024.

[10] J. Zhang, T. Fu, Y. Wang, J. Li, D. Xiao, J. Fan, et al., "An Alternately Optimized Generative Adversarial Network with Texture and Content Constraints for Deformable Registration of 3D Ultrasound Images," Physics in Medicine and Biology, vol. 68, no. 14, article no. 145006, July 2023.

[11] T. Murabayashi, R. Matsushima, and S. Sugimoto, "Novel Technique of Additional Anchor Plastic Stent Placement during Endoscopic Ultrasound-Guided Gallbladder Drainage," Endoscopy, vol. 56, no. 1, pp. 283-284, December 2024.

[12] T. Okuzono and K. I. Miyamoto, "Novel Anchoring Device for Endoscopic Ultrasound-Guided Gallbladder Drainage: Secondary Publication," Journal of Hepato-Biliary-Pancreatic Sciences, vol. 29, no. 7, pp. 825-831, July 2022.

[13] K. Iwano, T. Hayashi, and A. Katanuma, "Successful Re-Intervention of Endoscopic Ultrasound-Guided Gallbladder Drainage Using a Dual-Channel Endoscope," Digestive Endoscopy, vol. 34, no. 6, pp. 143-144, September 2022.

[14] K. G. Foley, M. J. Lahaye, R. F. Thoeni, M. Soltes, C. Dewhurst, S. T. Barbu, et al., "Management and Follow-Up of Gallbladder Polyps: Updated Joint Guidelines Between the ESGAR, EAES, EFISDS and ESGE," European Radiology, vol. 32, no. 5, pp. 3358-3368, May 2022.

[15] S. W. van der Merwe, R. L. J. van Wanrooij, M. Bronswijk, S. Everett, S. Lakhtakia, M. Rimbas, et al., "Therapeutic Endoscopic Ultrasound: European Society of Gastrointestinal Endoscopy (ESGE) Guideline," Endoscopy, vol. 54, no. 2, pp. 185-205, December 2022.

[16] A. Kamaya, C. Fung, J. L. Szpakowski, D. T. Fetzer, A. J. Walsh, Y. Alimi, et al., "Management of Incidentally Detected Gallbladder Polyps: Society of Radiologists in Ultrasound Consensus Conference Recommendations," Radiology, vol. 305, no. 2, pp. 277-289, November 2022.

[17] T. Kim, Y. H. Choi, J. H. Choi, S. H. Lee, S. Lee, and I. S. Lee, "Gallbladder Polyp Classification in Ultrasound Images Using an Ensemble Convolutional Neural Network Model," Journal of Clinical Medicine, vol. 10, no. 16, article no. 3585, August 2021.

[18] S. T. Hsu, Y. J. Su, C. H. Hung, M. J. Chen, C. H. Lu, and C. E. Kuo, "Automatic Ovarian Tumors Recognition System Based on Ensemble Convolutional Neural Network with Ultrasound Imaging," BMC Medical Informatics and Decision Making, vol. 22, article no. 298, 2022.