

# Unsupervised Anomaly Detection for Liver Lesions Identification Using Computed Tomography Imaging: A SimpleNet Framework Approach

Chen Run 24520241154760\*, information school class  
Jiang Xiaokang 31520241154488\*, information school class  
Li Linghang 24520240157768\*, information school class  
Li Lingjun 23020241154351\*, information school class  
Zhou Hongbo 23020241154365\*, information school class

\*School of Informatics, Xiamen University  
{chenrun, jiangxiaokang, lilinghang, lilingjun, zhouhongbo}@stu.xmu.edu.cn

## Abstract

This study presents an unsupervised anomaly detection framework for the early and accurate identification of liver lesions, including hepatocellular carcinoma (HCC), utilizing computed tomography (CT) imaging data. Traditional diagnostic methods, reliant on radiologists' expertise, suffer from inter-observer variability. Our approach leverages the SimpleNet framework, which employs pre-trained feature extraction, domain-specific adaptation, and anomaly feature generation to efficiently localize anomalies while addressing domain bias. We focus on adapting features to the target domain, generating synthetic anomalies, and utilizing clustering to differentiate normal and abnormal liver tissues. The Liver Tumor Segmentation (LiTS) dataset serves as a robust platform for developing and evaluating our proposed method. Our model achieves an Instance AUROC of 0.8823 and a Full-Pixel AUROC of 0.7984, demonstrating its potential to enhance automated diagnostic tools for HCC by detecting liver lesions with high accuracy and reliability. The integration of unsupervised clustering with advanced feature adaptation methods shows promise in improving the detection of subtle liver abnormalities and diagnostic accuracy.

## Introduction

The liver is a vital organ responsible for multiple critical functions, including metabolism, detoxification, protein synthesis, and bile production. Despite its regenerative capacity, the liver is vulnerable to chronic injuries, which may progress to cirrhosis and ultimately hepatocellular carcinoma (HCC), the most common primary liver malignancy (Christ and Parvathi 2011). HCC accounts for over 80% of liver cancer cases worldwide and is a leading cause of cancer-related mortality. Accurate and early detection of liver lesions, including HCC, is crucial for improving patient outcomes.

Medical imaging, particularly computed tomography (CT), is essential for the diagnosis, staging, and monitoring of liver diseases. CT imaging offers high-resolution multi-phase scans that provide detailed insights into the vascular and tissue characteristics of liver lesions (Dietrich, Dong, and Wang 2021). However, traditional diagnostic methods

heavily rely on radiologists' expertise, which may introduce inter-observer variability. In this context, unsupervised machine learning and anomaly detection techniques are emerging as promising solutions for automating diagnostic processes (Litjens et al. 2017).

Unsupervised anomaly detection methods aim to identify deviations from normal patterns in data without relying on labeled abnormal samples. Among these, clustering-based approaches group data into distinct clusters, enabling the identification of anomalies as outliers or underrepresented patterns. The SimpleNet framework exemplifies a state-of-the-art method for anomaly detection by leveraging pre-trained feature extraction, domain-specific adaptation, and anomaly feature generation (Liu et al. 2023). SimpleNet combines the strengths of synthesizing-based and embedding-based approaches to efficiently localize anomalies while addressing domain bias.

The Liver Tumor Segmentation (LiTS) dataset provides a robust platform for developing and evaluating liver tumor detection methods (Bilic and Christ 2023). By incorporating annotated CT scans from diverse clinical cases, it supports the validation of algorithms aimed at liver and tumor segmentation. The integration of unsupervised clustering with advanced feature adaptation methods offers the potential to enhance the detection of subtle liver abnormalities and improve diagnostic accuracy.

In this study, we propose to apply unsupervised clustering techniques inspired by the SimpleNet framework to CT imaging data from the LiTS dataset. Our approach focuses on adapting features to the target domain, generating synthetic anomalies, and utilizing clustering to differentiate normal and abnormal liver tissues. This methodology aims to advance automated diagnostic tools for HCC by achieving high accuracy and reliability in detecting liver lesions.

## Related Work

In recent years, researchers have proposed various deep learning models for liver tumor segmentation tasks on the LiTS dataset (Manjunath and Yashaswini 2024).

**Improved U-Net** model enhances attention to important regions by incorporating attention mechanisms into the skip connections of the U-Net, improving the segmentation ac-

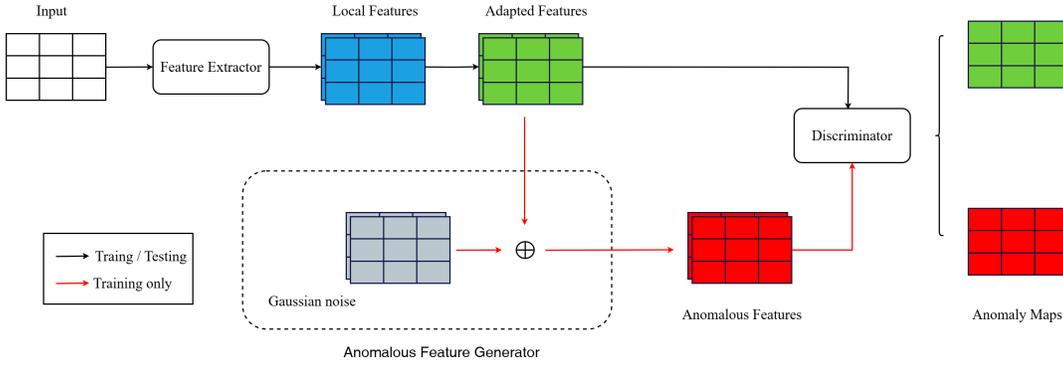


Figure 1: This is the structure of SimpleNet. As is shown to us, images will be processed by feature extractor to get their local features; Feature adaptor is used for fine-tuning them to adapted features; During training, anomalous feature generator will add some Gaussian noise to adapted features for anomalous features; at last, discriminator is trained by adapted features and anomalous features to draw the anomaly maps, or just give the anomaly maps of the adapted features while testing.

curacy of small tumors, with a Dice similarity coefficient reaching 0.69.

**MA-cGAN** is a model that combines multi-axis attention with conditional generative adversarial networks, capable of extracting key features from 3D CT images and generating high-quality segmentation results, achieving a Dice coefficient of 96.95 for liver segmentation and 78.53 for tumor segmentation.

**YOLOv8** framework was developed for lightweight liver and tumor segmentation, balancing computational cost and accuracy, making it suitable for clinical practice.

**G-Unet** model, which was evaluated on the LiTS dataset with various configurations, demonstrating a Dice global score of up to 72.9%, proving its effectiveness in liver tumor segmentation.

**DeepLabv3+**, which utilize dilated convolutions, have also been applied to this task, enhancing feature extraction capabilities and further improving segmentation accuracy (Tanfoni et al. 2024).

The research on these various models showcases the wide application and potential of deep learning techniques in the automated segmentation of liver tumors while also highlighting the challenges faced in handling complex medical imaging tasks.

This diverse range of model approaches provides a wealth of options for achieving efficient and accurate liver tumor segmentation and fosters continued development in this field.

## Method

Our work uses SimpleNet, which is designed as a simple and efficient model in Industrial Anomaly Detection(IAD), with the following key features:

- Utilizing a pre-trained feature extractor to extract local features.
- Introducing a feature adaptor to fine-tune the features from the pre-trained features extractor to the target domain, minimizing domain bias.

- Generating synthetic anomalous features by incorporating Gaussian noise to the normal feature space.
- Employing a simple binary classifier as anomaly discriminator to differentiate normal features from anomalous features.
- During the inference phase, the synthetic anomalous feature generator is discarded, ensuring fast and efficient network performance.

### Feature Extractor

This module intends to provide high-quality features for subsequent modules. Local features extracted from this part is referred to this paper (Roth et al. 2022). For every image  $x_i \in \mathbb{R}^{H \times W \times 3}$  in training set  $\mathcal{X}_{train}$  and test set  $\mathcal{X}_{test}$ , the pre-trained network  $\phi$  (such as WideResNet50) will extract features from all hierarchies, which actually is unnecessary. Considering the domain bias between the pre-trained dataset and the target dataset, SimpleNet chooses the features from a subset of levels, defined as  $L$ . In level  $l$ , we can get feature map as  $\phi^{l,i} \sim \phi^l(x_i) \in \mathbb{R}^{H_l \times W_l \times C_l}$ . At location  $(h, w)$ , we define its neighborhood as

$$\mathcal{N}_p^{(h,w)} = \{(h', y') | h' \in [h - \lfloor p/2 \rfloor, \dots, h + \lfloor p/2 \rfloor], y' \in [w - \lfloor p/2 \rfloor, \dots, w + \lfloor p/2 \rfloor]\} \quad (1)$$

By using adaptive average pooling as aggregation function  $f_{agg}$ , local feature  $z_{h,w}^{l,i}$  can be obtained by aggregating the features within the neighborhood.

$$z_{h,w}^{l,i} = f_{agg}(\{\phi_{h',y'}^{l,i} | (h', y') \in \mathcal{N}_p^{h,w}\}) \quad (2)$$

In order to integrate information from different hierarchies, this module resizes all related feature maps to the same size  $(H_0, W_0)$ , then we get the feature map  $o^i \in \mathbb{R}^{H_0 \times W_0 \times C}$  by following process:

$$o^i = f_{cat}(resize(z^{l,i}, (H_0, W_0)) | l \in L) \quad (3)$$

At last, just simplify the process mentioned above and we can define the feature extractor  $F_\phi$  as follows:

$$o^i = F_\phi(x^i) \quad (4)$$

## Feature Adaptor

Receiving those local features from feature extractor, this module transfer them into adapted features, reducing the bias between the distribution of ImageNet and that of target domain. The feature adaptor  $G_\theta$  is defined by projecting the local feature  $o_{h,w}$  to the adapted feature  $q_{h,w}$  as shown in follows:

$$q_{h,w}^i = G_\theta(o_{h,w}^i) \quad (5)$$

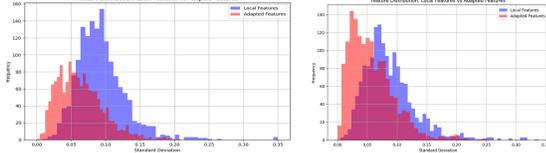


Figure 2: As is shown in this graph, adapted features have much more compact space than local, which makes the training more efficient.

## Anomalous Feature Generator

While training, this module generates forged anomalous features based on those adapted features, providing negative samples for discriminator. To train a binary classifier, we need to prepare normal and abnormal samples. Since this is an unsupervised model, we can only get normal features directly from input images, therefore, this generator helps to generate some anomalous features. By adding Gaussian noise  $\epsilon$  on normal features  $q_{h,w}^i \in \mathbb{R}^C$ , the anomalous feature  $q_{h,w}^{i-}$  can be expressed as follows:

$$q_{h,w}^{i-} = q_{h,w}^i + \epsilon \quad (6)$$

## Discriminator

As the last module, discriminator transfers its input features to anomaly score map, which enables pixel-wise anomaly detection. Discriminator does nothing but estimating the normality at each location. After training the discriminator  $D_\Psi$ , we can predict the anomaly score map by calculating  $D_\Psi(q_{h,w})$

## Loss

Since we use a binary classifier as the discriminator, we just need to ensure the positive being positive and the negative being negative, therefore, we choose a simple  $l1$  loss as following:

$$l_{h,w}^i = \max(0, th^+ - D_\Psi(q_{h,w}^i)) + \max(0, -th^- + D_\Psi(q_{h,w}^{i-})) \quad (7)$$

$th^+$  and  $th^-$  are used to prevent overfitting. And the objective of training is

$$\mathcal{L} = \min_{\theta, \Psi} \sum_{x^i \in \mathcal{X}_{train}} \sum_{h,w} \frac{l_{h,w}^i}{H_0 \times W_0} \quad (8)$$

## Scoring Function

The anomaly score is predicted by discriminator as

$$s_{h,w}^i = -D_\Psi(q_{h,w}^i) \quad (9)$$

Accordingly, this allows us to compute the performance metrics for both pixel-wise and image-wise evaluations.

$$S_{AL}(x_i) := s_{h,w}^i | (h,w) \in H_0 \times W_0 \quad (10)$$

$$S_{AD}(x_i) := \max_{(h,w) \in H_0 \times W_0} s_{h,w}^i \quad (11)$$

## Experiments

### Dataset Description

In this work, we utilized the Liver Tumor Segmentation Challenge (LiTS) dataset. It is a benchmark created under the purviews of ISBI 2017 and MICCAI 2017 intended for fostering improvement in automatic liver lesion segmentation. The dataset comprises contrast-enhanced abdominal CT scans along with segmentation annotations from different clinical sites worldwide. It consists of 130 CT scans for training, as well as an additional 70 for testing.

### Evaluation Metrics

We evaluate our approach using the following three metrics, tailored for unsupervised anomaly detection and localization:

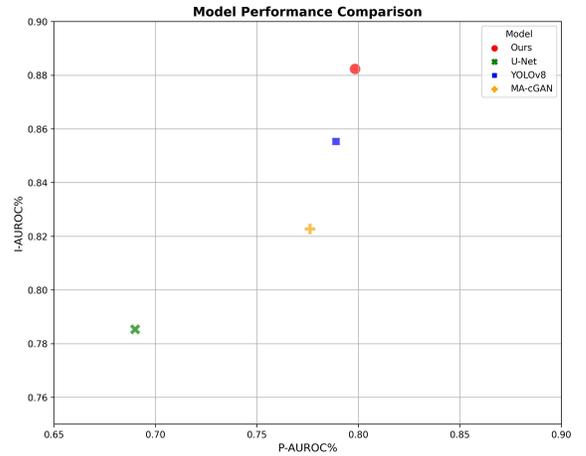


Figure 3: Performance of different on database, our model behave the best on both I-AUROC% and P-AUROC%.

- Instance AUROC quantifies the model's ability to differentiate between anomalous and normal images by computing the area under the receiver operating characteristic curve at the instance level.
- Full-Pixel AUROC evaluates pixel-level anomaly detection across the entire image by measuring the AUROC for all pixels, providing an overall assessment of anomaly localization performance.
- Anomaly-Pixel AUROC focused on detecting anomalous regions, this metric calculates the AUROC for pixels within predefined regions of interest, offering finer insight into the model's ability to localize specific anomalies.

model	I-AUROC%	P-AUROC%
Ours	<b>0.8823</b>	<b>0.7984</b>
U-Net	0.7853	0.6900
YOLOv8	0.8553	0.7890
MA-cGAN	0.8227	0.7762

Table 1: Comparison of SimpleNet with state-of-the-arts works on database. Image-wise AUROC (I-AUROC) and pixel-wise AUROC(P-AUROC) are displayed in each entry as I-AUROC%/P-AUROC%

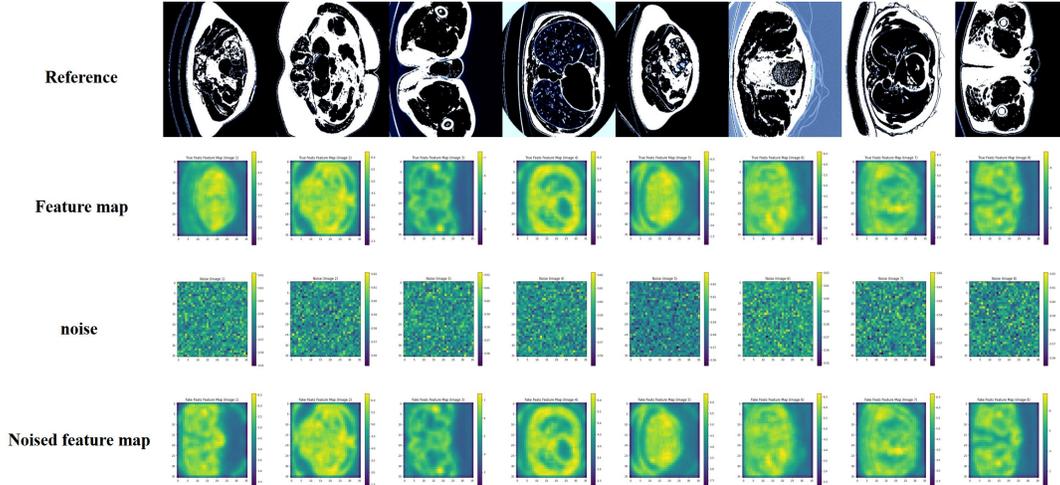


Figure 4: Compare the extracted feature maps from normal images with those extracted from noisy images. This comparison illustrates how well the model can segment out anomalies.

## Results

The performance of the proposed framework is illustrated through various experiments. From the results obtained, the model gives an **Instance AUROC of 0.8823**, which shows its capability to discern between normal and anomalous images. It attains a **Full-Pixel AUROC of 0.7984**; therefore, proving that it can identify anomalies at the pixel level across entire images. However the **Anomaly-Pixel AUROC is 0.4732**, proving somewhat weaker ability to spot and less salient anomalies at that finer level.

The feature maps for normal images form the base representation of healthy patterns and structures. By introducing noise to these features, the model simulates realistic uncertainty and reconstructs the features to evaluate its robustness and precision. The discrepancies observed between the original and noised feature maps indicate regions where the model struggles to accurately recreate specific patterns. These areas often correspond to potential anomalies, as they deviate from the learned normal patterns. This systematic comparison demonstrates the model’s ability to handle noise-augmented scenarios while retaining sensitivity to structural deviations, making it well-suited for unsupervised anomaly detection tasks.

## Conclusion

This study showcases the utility of an unsupervised anomaly detection framework for identifying liver lesions. By lever-

aging feature extraction, noise injection, and reconstruction techniques, the model efficiently highlights deviations indicative of anomalies. The approach demonstrates strong potential to assist in automating diagnostic workflows, reducing dependency on manual interpretation. However, challenges remain in achieving precise localization of subtle anomalies, which will guide future improvements to ensure the framework’s broader clinical applicability.

## Acknowledgments

The authors would like to express their gratitude for the support and resources provided during the course of this study. We acknowledge the contribution of the Liver Tumor Segmentation (LiTS) dataset, which facilitated the development and evaluation of our liver tumor detection methods. Special thanks are due to the clinical sites worldwide for their involvement in creating the benchmark dataset for automatic liver lesion segmentation. We also appreciate the insights and discussions from our friends and classmates, which significantly contributed to the refinement of our unsupervised anomaly detection framework. Lastly, we thank the reviewers especially for the teacher and teaching assistant. Their constructive comments and suggestions have helped us to improve the quality of this manuscript.

## References

- Bilic, P.; and Christ, P. 2023. The Liver Tumor Segmentation Benchmark (LiTS). *Medical Image Analysis*, 84: 102680.
- Christ, M. C. J.; and Parvathi, R. M. S. 2011. Fuzzy c-means algorithm for medical image segmentation. In *2011 3rd International Conference on Electronics Computer Technology*, volume 4, 33–36.
- Dietrich, C. F.; Dong, Y.; and Wang, W.-P. 2021. *Contrast Enhanced Ultrasound: How to Perform It in Liver Tumors?*, 15–24. Singapore: Springer Singapore. ISBN 978-981-16-1761-4.
- Litjens, G.; Kooi, T.; Bejnordi, B. E.; Setio, A. A. A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J. A.; van Ginneken, B.; and Sánchez, C. I. 2017. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42: 60–88.
- Liu, Z.; Zhou, Y.; Xu, Y.; and Wang, Z. 2023. SimpleNet: A Simple Network for Image Anomaly Detection and Localization. arXiv:2303.15140.
- Manjunath, R.; and Yashaswini. 2024. Automated segmentation of liver tumors from computed tomographic scans. *Journal of Liver Transplantation*, 15: 100232.
- Roth, K.; Pemula, L.; Zepeda, J.; Schölkopf, B.; Brox, T.; and Gehler, P. 2022. Towards Total Recall in Industrial Anomaly Detection. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 14298–14308.
- Tanfoni, M.; Ceroni, E. G.; Maggini, M.; Pancino, N.; and Bianchini, M. 2024. A Hybrid Deep Learning Approach for Liver Tumor Segmentation Using DeepLabV3+ and Hidden Markov Models. In *2024 IEEE International Symposium on Systems Engineering (ISSE)*, 1–5.