COMPARATIVE ANALYSIS OF LANET AND FE-NET IN IMAGE SEGMENTATION

Zhao Di¹, Alevtina Gourinovitch¹, Shadiye Sultanova²,

Baloghlan Najafov², Gizilgul Israfilova²

¹Belarussian State University, Minsk, Belarus ²Azerbaijan Technical University, Baku, Azerbaijan 3189124246@qq.com, alevt5555@gmail.com, shadiye.sultanova@aztu.edu.az, baloglan.necefov@aztu.edu.az, qizilgul.israfilova@aztu.edu.az

Abstract

Image segmentation in infocommunications helps analyze infrastructure, enhance video quality, manage networks to improve efficiency and service quality. Biomedical image segmentation is very important for the medical imaging. It requires precise delineation of anatomical structures and pathological regions for various diagnostic tasks (for example, tumor detection and treatment planning) in improving clinical outcomes. In recent decades, the automatic medical segmentation methods which are based on deep learning (DL) models, such as convolutional neural network (CNN) architectures, have facilitated the automatic delineation of organ and lesion boundaries. This changed traditional manual segmentation approaches and improved efficiency and accuracy in clinical practice. This paper presents a comparative analysis of two DL-based models: LANet and FE-Net. LANet employs an Efficient Fusion Attention (EFA) module and an Adaptive Feature Fusion Decoder (AFF) module to improve segmentation efficiency and precision. In contrast, FE-Net integrates a Feature Awareness Module (FAM) for enhances its features to capture multi-scale and process segmentation details. The advantages of both methods in handling different scales, details, and edges were investigated and experimentally evaluated on various public datasets, and their performance in specific scenarios was assessed. The experiment showed that LANet is superior in computational efficiency and feature refinement. FE-Net shows superior performance in handling complex variations and edge details. The source code of LANet and FE-Net can be found on GitHub at [https://github.com/tyjcbzd/LANet] and [https://github.com/tyjcbzd/FE-Net].

Keywords: Image segmentation, service quality, deep learning, FE-Net, LANet.

I. Introduction

Biomedical image segmentation is crucial for various medical applications, including clinical diagnosis, treatment, and quantitative analysis, being indispensable for both clinical analysis and surgical procedures. This method involves partitioning images into multiple regions or Regions of Interest (ROIs) based on specific characteristics such as color, texture, and shape, and subsequently extracting these areas for further analysis. Accurate segmentation is vital for ensuring the quality of clinical diagnoses and treatment plans. Traditional segmentation methods, including threshold-

based approaches and region-growing algorithms, often depend on manual or semi-automatic processes [1]. These methods are not only time-consuming and labor-intensive but also susceptible to the operator's experience and subjective judgment, which limits their widespread applicability and accuracy in clinical settings. However, with the advent of deep learning technologies, the field of biomedical image segmentation has experienced a transformative shift from conventional manual methods to automated, algorithm-driven approaches that deliver higher precision and efficiency.

In recent years, the rapid advancement of deep learning techniques, particularly Convolutional Neural Networks (CNNs) with their robust feature representation capabilities, has significantly improved performance in biomedical image segmentation. U-Net, a prominent algorithm for biomedical image segmentation based on CNNs, introduced skip connections to merge multi-level features, achieving exceptional segmentation results. Numerous U-Net-based variant architectures, such as Res-UNet[2], U-Net++[3], and TransUNet [4], have been proposed for biomedical image segmentation. These models can automatically delineate organ or lesion contours, effectively overcoming the limitations of manual segmentation. The rapid development of computer hardware, coupled with advancements in deep learning, has enabled these models to demonstrate remarkable performance in the automatic segmentation of objects of interest. However, biomedical images frequently affected by noise, exhibit complex backgrounds, and display similar appearances across different tissues, which complicates the extraction of valuable information. The traditional CNN methods may face challenges in effectively balancing the capture of relevant information with the elimination of redundant features, which can lead to poor segmentation performance and inefficient feature representation. Furthermore, the fusion of low-level and high-level features poses a significant challenge for biomedical image segmentation. This kind of fusion process requires adaptability and context awareness to ensure the seamless fusion of complementary information to improve the network's ability to accurately analyze and represent the input features.

In this context, our previous efforts created two state-of-the-art (SOTA) biomedical image segmentation methods: LANet [5] and FE-Net[6]. LANet consists of an Efficient Fusion Attention (EFA) module and an Adaptive Feature Fusion (AFF) decoding block, aiming to improve the abilities of feature extraction by effectively capturing task-specific information and minimizing redundancy in channel and spatial dimensions [5]. Conversely, FE-Net introduces a Feature-Aware Module (FAM) that emphasizes important features while suppressing irrelevant ones, utilizing an encoder-decoder architecture to address multi-scale challenges in biomedical images [6]. In this paper evaluates the performance of these two models using three public datasets: Kvasir-SEG, low-grade glioma (LGG), and the 2018 Data Science Bowl, applying standard evaluation metrics including precision, Dice coefficient, mean Intersection over Union (mIoU) and recall.

II. Methods Overview Data Description and Preprocessing

The three public datasets were used in this article: Kvasir-SEG [7], LGG [8], and the 2018 Data Science Bow I[9]. These datasets contain a variety of medical images and multiple pathological conditions. They can provide a comprehensive testing environment to evaluate the generalization abilities and robustness of the algorithms. Below is the detailed description of each dataset:

Kvasir-SEG: The dataset includes 1,000 gastrointestinal polyp images of diverse resolutions, the features segmentation masks of the database is annotated by knowledgeable gastroenterologists. These notations contain regular tissue, polyps, ulcers, and various other lesions, offering a valuable resource for segmentation purposes.

LGG: This collection comprises 1,310 RGB medical images, each having a resolution of 256×256 pixels. Acquired from the Cancer Imaging Archive at the National Cancer Institute, this compilation highlights LGG studies and showcases FLAIR mode photographs from individuals diagnosed with

TCGA. Neuroradiology experts have manually labeled irregular sections in these scans to aid in the development of automated segmentation techniques.

2018 Data Science Bowl: This collection comprises 670 cellular images, each featuring a 256×256 pixels clarity, showcasing a varied spectrum of cellular information, encompassing both cancerous and healthy cells.

The three public datasets were divided into training, validation, and testing segments by employing random selection in an 8:1:1 ratio. The goal of this segmentation approach is to ensure the model receives enough samples in training and to assess its ability for generalization and predictive precision through validation and test datasets. Comprehensive details on each dataset are provided in Table 1.

Dataset	Images	Size	Train	Validation	Test	Application
Kvasir-SEG	1000	Variables	800	100	100	Colonoscopy
LGG	1310	256 x 256	1048	131	131	Brain
2018 Data Science Bowl	670	256 x 256	530	67	67	Nuclei

Table 1: The information of the dataset

In the phase of data preprocessing, as shown in Figure 1, the article standardized all images to maintain data consistency. The process entailed adjusting the image scale to fit the model's input needs and conducting normalization to reduce discrepancies due to various devices and imaging scenarios. In order to enhance the model's robust and improve its adaptability to unseen samples, the article implemented multiple techniques for data augmentation applied to the training data. These methods utilized included random cropping and rotation, which help to introduce variability and enhance training set's diversity, thus improving the model to generalize better to new data.

However, in order to maintain the integrity and fairness of the test results, the article refrained from applying any data augmentation to the test set. This approach ensures that the evaluation metrics reflect the model's true performance on unaltered and real-world data.



Figure 1: Data augmentation techniques applied on the LGG dataset

III. Model Architecture

LANet structure

As shown in Figure 2, LANet is based on a lightweight MobileViT backbone network, incorporating the EFA module and the AFF decoding module. The EFA module extracts key task-relevant features and reduces redundancy through channel and spatial attention. Meanwhile, the AFF module enhances feature representation by combining low-level and high-level features from the encoding path during the decoding process. The lightweight design of LANet makes it suitable for resource-constrained environments, such as real-time clinical devices. The LANet source code is hosted on GitHub at [https://github.com/tyjcbzd/LANet].



Figure 2: The structure of LANet *a*- LANet architecture overview; *b*- EFA Block structure; *c*- AFF-D Block structure.

FE-Net structure



Figure 3: The architecture of the proposed network and module: *a* - *FE*-Net architecture; *b* - feature-aware module flowchart

As illustrated in Fig.3, FE-Net utilizes a traditional encoder-decoder architecture, augmented by the incorporation of a FAM aimed at improving detail segmentation. The FAM module utilizes attention mechanisms to highlight important features and reduce noise during feature extraction, and it also uses skip connections to facilitate the fusion of low-level and high-level features. In contrast to other segmentation models, FE-Net uses bilinear interpolation instead of transposed convolution during the decoding phase. This strategic choice not only reduces computational complexity but also preserves the smoothness of the resulting images. The FE-Net source code is hosted on GitHub at [https://github.com/tyjcbzd/FE-Net].

IV. Evaluation Metrics

To assess the effectiveness of LANet and FE-Net, this study employs a range of standard evaluation metrics, including precision, the Dice coefficient (commonly referred to as the F1 score), recall, and mIoU. The Dice coefficient serves as a measure of the accuracy of the segmentation results, while the Intersection over Union (IoU) quantifies the degree of overlap between the model's predictions and the ground truth annotations. Precision reflects the model's accuracy in identifying target objects, whereas recall indicates the model's effectiveness in detecting all relevant targets. All evaluation metrics are derived from true positives (TP) and false positives (FP).

1. Precision: Precision measures the proportion of TP in the predicted segmentation results, in particular the ratio of actual positives among the samples predicted as positive. A high precision indicates that there are fewer FP in the model's predictions. Precision is calculated using the following formula:

$$Precision = \frac{TP}{TP + FP}$$
(1)

2. Dice Coefficient (F1 Score): The Dice score measures the accuracy by comparing the overlap region between the predicted segmentation results and the ground truth annotations. It quantifies the model's ability to match the true segmentation in biomedical image analysis tasks. The score reflects the extent of overlap, with scores between 0 and 1. The higher scores indicate that a closer match between the model's output and the actual labels. The Dice coefficient is computed with the following formula:

Dice =
$$\frac{2TP}{2TP + FP + FN}$$
 (2)

3. Recall: Recall measures the ratio of true positive samples that the model successfully detected, specifically the ratio of correctly detected true positives among all true positive samples. A high recall rate indicates that the model can identify more true positive samples, thereby reducing the number of false negatives (FN). Recall can be calculated using the following formula:

$$\operatorname{Recall} = \frac{TP}{TP + FN} \tag{3}$$

4. MIoU: MIoU is a metric used to evaluate the performance of image segmentation tasks.It evaluate the degree of overlap between the predicted results from a model and the true labels by calculating the ratio of the intersection to the union for each category and then averaging this ratio across all categories. MIoU is commonly used in multi-class segmentation tasks as a comprehensive performance metric. Assuming there is a total of k categories, the calculation formula is as follows:

$$mIOU = \frac{1}{k} \sum_{i=0}^{k} \frac{TP}{TP + FP + FN}$$
(4)

V. Experiment results overview Quantitative comparison

Table 2: Comparison of experimental results on three different medical datasets namely LGG, Kvasir-SEG and 2018
Data Science Bowl. The best results are in bold.

Dataset	Module	mIoU	Dice	Precision	Recall
LGG	LANet	0.854	0.906	0.926	0.866
	FE-Net	0.914	0.963	0.988	0.962
Kvasir-SEG	LANet	0.851	0.911	0.949	0.903
	FE-Net	0.830	0.889	0.915	0.901
2018 Data Science	LANet	0.871	0.930	0.946	0.918
Bowl	FE-Net	0.864	0.924	0.986	0.939

1. Performance on the Kvasir-SEG Dataset: On the Kvasir-SEG dataset, the performance difference between LANet and FE-Net decreased. LANet achieved a mIoU of 0.851, while FE-Net scored 0.830, indicating that both of the two models performed comparably, with LANet having a slight edge. However, in terms of precision, LANet outperformed FE-Net with a value of 0.949, which underscores higher accuracy of LANet in identifying positive samples. In terms of recall, FE-Net slightly surpassed LANet with a value of 0.903 compared to the value of 0.901 of LANet. It is clear that FE-Net had a slight advantage in the recall of detecting positive samples.

2. Performance on the LGG Dataset: LANet achieved a mIoU value of 0.854 On this dataset, while FE-Net exceeded this with an mIoU value of 0.914. This result indicates that FE-Net is more effective in identifying and segmenting lesion areas. Regarding Dice coefficient, precision, and recall, FE-Net also outperformed LANet, achieving values of 0.963, 0.988, and 0.962, respectively, compared to the values of 0.906, 0.926, and 0.866 of LANet. The numerous benefits of FE-Net across various metrics further confirm its dominance in precise segmentation.

3. Performance on the 2018 Data Science Bowl Dataset: On the dataset, the performance of LANet and FE-Net diverged from the trends observed in the previous datasets. LANet achieved a mIoU of 0.871, while FE-Net scored 0.864, which can indicate a minimal variance. However, regarding the Dice coefficient, accuracy, and recall, FE-Net outperformed LANet with values of 0.924, 0.986, and 0.939. In comparison, LANet scored 0.930, 0.946, and 0.918. These results demonstrate that FE-Net's strengths in these metrics reflect its adeptness in controlling segmentation details and boundaries.

VI. Qualitative Analysis

Figure 4 shows a comparison of the segmentation results of LANet and FE-Net on different datasets. The unique strengths of each model are evident. LANet is particularly effective at segmenting larger targets. FE-Net excels at handling details and edges. For example, in the segmentation of gastrointestinal polyp images from the Kvasir-SEG dataset, LANet achieved impressive boundary alignment, resulting in a precision value of 0.949. On the other hand, FE-Net demonstrated a greater ability to capture finer details, achieving a recall value of 0.915, which highlights its sensitivity to subtle features. When processing cell images with complex boundaries, FE-Net was able to segment these boundaries more accurately, while LANet showed higher overall

accuracy in identifying larger targets. LANet excelled in detecting larger tumors, in contrast to FE-Net, who was more adept at detecting smaller or initial stage tumors.

LANet and FE-Net demonstrated good adaptability in various imaging techniques. However, FE-Net outperformed LANet in processing high-contrast images. Additionally, LANet exhibited robust in dealing with noise, particularly with noise patterns commonly found in MRI images. Despite FE-Net is more sensitive to noise in some situations, its feature perception module can reduce noise impact by learning contextual information from the medical images.



Figure 3: The comparison of the segmentation results of LANet and FE-Net

VII. Conclusion

The study was conducted jointly with the Department of Magnetic Resonance Imaging of the N. N. Alexandrov National Oncology Center.

It provides a comparative analysis of LANet and FE-Net in biomedical image segmentation. This study highlights joint efforts to improve biomedical image analysis capabilities. Both models exhibit unique strengths and limitations. This significantly impacts their performance and applicability in practice. LANet and FE-Net showed excellent performance and unique strengths in biomedical image segmentation. However, they also demonstrated limitations that impact their performance and applicability.

LANet's lightweight architecture, augmented with EFA and AFF modules, optimizes computational efficiency and processing speed, making it suitable for resource-constrained settings such as mobile devices and real-time clinical applications. Accuracy may not be achieved in high-resolution segmentation tasks and datasets with significant variability and noise.

FE-Net with FAM excels in capturing multi-scale features and segmentation detail. This is especially true for images with complex boundaries and fine structures. The inclusion of bilinear interpolation in FE-Net reduces the computational burden while maintaining image smoothness, which is an important factor for real-time applications. However, the complexity of FAM increases the computational requirements, potentially hindering its deployment in resource-constrained environments. Moreover, FE-Net's sensitivity to noise poses challenges in noisy imaging modalities such as low-dose CT or MRI. Its detailed feature extraction increases the risk of overfitting when trained on limited datasets.

Both methods show significant promise for the task of biomedical image segmentation. Future research can aim to integrate their respective strengths to obtain more efficient and accurate segmentation results.

References

[1] X. Liu, Z. Deng, Y. Yang, Recent progress in semantic image segmentation, Artif. Intell. Rev. 52 (2019) 1089-1106.

[2] X. Xiao, S. Lian, Z. Luo, et al., Weighted Res-UNet for high-quality retina vessel segmentation, 2018 9th International Conference on Information Technology in Medicine and Education (ITME), IEEE, (2018) 327-331.

[3] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh,, & J. Liang, Unet++: A nested u-net architecture for medical image segmentation, Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20,2018, Proceedings 4.Springer International Publishing,(2018) 3-11.

[4] J. Chen, Y. Lu, Q. Yu, et al., TransUNet: Transformers make strong encoders for medical image segmentation, arXiv preprint arXiv:2102.04306, (2021).

[5] Y. Tang, D. Pertsau, D. Zhao, D. Kupryianava, M. Tatur, LANet: Lightweight Attention Network for Medical Image Segmentation, Information Technologies and Their Applications. ITTA 2024. Communications in Computer and Information Science, Springer, Cham.(2025) 213-227.https://doi.org/10.1007/978-3-031-73420-5_18

[6] D. Zhao., Y. Tang, A.B. Gourinovitch, Effective Algorithm for Biomedical Image Segmentation. Doklady BGUIR. (2024) 84-92. https://doi.org/10.35596/1729-7648-2024-22-3-84-92.

[7] D. Jha, P.H. Smedsrud, M.A. Riegler, et al., Kvasir-SEG: A segmented polyp dataset, Multi Media Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020, Proceedings, Part II 26, Springer International Publishing, (2020) 451-462.

[8] M. Buda, A. Saha, M.A. Mazurowski, Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm, Comput. Biol. Med. 109 (2019) 218–225.

[9] J. C. Caicedo, A. Goodman, K. W. Karhohs, B. A. Cimini, J. Ackerman, M. Haghighi, A. E. Carpenter, Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. Nature methods,(2019) 1247-1253.